

MEDLINE REFERENCE	PUBMED ID
20530913	11076861
4 (sites)	
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.	
TITLE JOURNAL REFERENCE AUTHORS	Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001) 5 (bases 1 to 2696) Adechi,J., Aizawa,K., Akahira,S., Akimura,T., Anno,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bulic,C., Carninci,P., Fukuda,S., Fukushima,Y., Furuno,M., Hangack,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Horii,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kao,H., Kawai,Y., Kojima,Y., Konno,H., Kouda,M., Komura,S., Kutlhara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schirral,L., Shibata,K., Shihara,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toyota,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y. Direct Submission Submitted (10-JUL-2000) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216) Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GGAAGAGAAGATCCCAAGACCTCTTTTCTTTTTTTTAA 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 7.5 and subtraction to Rot = 37.5. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAAGAGATTCGAGTGTAATTAAATTAAATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOFR.
FEATURES SOURCE	location/Qualifiers 1..2696 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="MGD:MGI:1896093" /db_xref="taxon:10090" /clone_lib="2810054N23" /cdonc_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="10, 11 days embryo" 112..1524 /gene="Traip" /locus="112..1524 /gene="Traip" /note="TRAF-interacting protein data source:MGD, source key:MGI:1096377, evidence:ISS putative" /codon_start=1 /protein_id="PAB28567.1" /db_xref="GI:12850019" translation="MPRALCTICSPDFHSRDVAIHCHTHFLQCLIMPEFAER TPCGRIGVGRKTITINKLFFDLAOEEENVDAELNELDSVAQLSKOREKRRDSA IIDTLRPTLEARNATVESILNALNAEMLCSTLKQMKLFLEORODETKOAREKHRLK CKMTAKROIILLSSQRSSEVEEMIRDMGVGSAAVOCLAVCYSKEINYLKERKRT GGADLMRLDKLVLSRSKLTKINTLELDALKELSAQCDLOSADOETSLKKKIILQG TSLSPATLESRLVRESPPAVPMKNRPRLHPGEFGEDLNTTFVNTPPTOTSQSG HCLPKKCLEPARSPMONVLKKHKHAKPSQSLLGGRCVGEGLDELGAAPLFTLN AVLGQKPNNFTAESRSTSVVRVGFDGGRFKFTIPRDTTIIRPVASKSKSKSR

Query Match	Best Local Similarity	Score	DB	Length
Matches 1939;	Conservative 0;	Mismatches 15;	Indels 4;	Gaps 3;
96.0%;	99.0%;	Pred. No. 0;		
polyA_signal	/note="putative"			
polyA_site	/note="putative"			
BASE COUNT	699 a	676 c	711 g	610 t
ORIGIN				
VRKTVSSASQPKDLFLCQ"				
2678..2683				
4 ACGAGTGGCGGTGACCGAAATTTGAAGAACCGGACGGTGGCGGTTCC--CACCAAC	61			
2 AAGTCGTGGGTGAGAGCAAAATTTGAAGAACCGGACGGTGGCGGTTCCCGCCACCAAC	61			
62 TGTGTCTGTCTCTGGGACGTGGTCCCTGGGCTGCTTGAAGTGAGGACCATATGCTATTC	121			
62 TGTGTCTGTCTGGGACGTGGTCCCTGGGCTGCTTGAAGTGAGGACCATATGCTATTC	121			
122 TCTCTGTGACATATCTCTCCGACTTCTTGATATCACTCCGTGACGCGGTGCTATTC	181			
122 GCGCTCTGTGCACTATCTCTCCGACTTCTTGATATCACTCCGTGACGCGGTGCTATTC	181			
182 ACTGTGGCCACACTTTCTATCTGCAATGCGCTAATCCAGTGGTTTGAGACAGCACCAGTC	241			
182 ACTGTGGCCACACTTTCTATCTGCAATGCGCTAATCCAGTGGTTTGAGACAGCACCAGTC	241			
242 GCACCTGCCACAGTGTGATATCCAGGTTGGCAAAAAGATATTTAAACAACCTTTTCT	301			
242 GCACCTGCCACAGTGTGATATCCAGGTTGGCAAAAAGATATTTAAACAACCTTTTCT	301			
302 TTGACCTGCCACAGGAAAGAGAGATGCTTGATGACGAATTTCTAAAGAATGAACGTGG	361			
302 TTGACCTGCCACAGGAAAGAGAGATGCTTGATGACGAATTTCTAAAGAATGAACGTGG	361			
362 ACAGGCTCAAAAGCTTCTCCAGAAAAGACAGGAGAAACGGGACACCGAGCCATTTA	421			
362 ACAGGCTCAAAAGCTTCTCCAGAAAAGACAGGAGAAACGGGACACCGAGCCATTTA	421			
422 TTGACACTTACCGGACACCTCGGAAGAAGCGCAATGCTACCGTGGAGTCCCTACGAACG	481			
422 TTGACACTTACCGGACACCTCGGAAGAAGCGCAATGCTACCGTGGAGTCCCTACGAACG	481			
482 CCTTAAACAAGGCAAGATGCTGTCCACCTGAAAACAGATGAAGATTTCTGGAGGC	541			
482 CCTTAAACAAGGCAAGATGCTGTGTCCACCTGAAAACAGATGAAGATTTCTGGAGGC	541			
542 AGCGGACAGATGAGACCAAAAGCTCGGAGAGAGGCCACCGACTCAAGTCAAGATGA	601			
542 AGCGGACAGATGAGACCAAAAGCTCGGAGAGAGGCCACCGACTCAAGTCAAGATGA	601			
602 AAACCATGAGCAAAATTGAGCTCTTACCTCCAGAGCCAGCGCTTCTGAGGTGAGAGATGA	661			
602 AAACCATGAGCAAAATTGAGCTCTTACCTCCAGAGCCAGCGCTTCTGAGGTGAGAGATGA	661			
662 TTTCGAGACATGGGTGTGGACAGTACGAGGCTGAGACACTGGCTGTGATCTGCTGCC	721			
662 TTTCGAGACATGGGTGTGGACAGTACGAGGCTGAGACACTGGCTGTGATCTGCTGCC	721			
722 TCAAAAGAGATGATGAGATCTGGAAGAAAGCTCGGAAGGCCACAGGGGAACCTGGTGACA	781			
722 TCAAAAGAGATGATGAGATCTGGAAGAAAGCTCGGAAGGCCACAGGGGAACCTGGTGACA	781			
782 GGTGGAAGAGATTTGGTGTCTCTAGAGCAAGTTGAAGACTCTCAACACTGAGCTGG	841			
782 GGTGGAAGAGATTTGGTGTCTCTAGAGCAAGTTGAAGACTCTCAACACTGAGCTGG	841			
842 ATCAGGCCAAGTTAGAACTGAGGTGACGCCAGAAAGACTTACAAAGTCTGACCAAGGAGA	901			
842 ATCAGGCCAAGTTAGAACTGAGGTGACGCCAGAAAGACTTACAAAGTCTGACCAAGGAGA	901			

QY	902	TCACGACCCTTAAGAAAGAAAGTCTGATGATGATCTTCACAGGAACTTTAGCCTGCCTCC	-GC6	960
Dp	902	TCACGAGACCCTAAGAAAAGAAAG	-CTGATGATGATCTTCACAGGAACTTTAGCCTGCCTCCGCG	960
QY	961	ACCAATGAGACGGTCACAGCCCGCTGGTTTTTTTGAGAGCCCGCCCGTGGAGATGATGAAAC		1020
Dp	961	ACCAATGAGACGGTCAAGCCCGCTGGTTTTTTTGAGAGCCCGCCCGTGGAGATGATGAAAC		1020
QY	1021	CCGAGGCTTCACACAGCCACCCTTGGGTGATGAGATTTGATCTCAATTAACACCTTTGATGTA		1080
Dp	1021	CCGAGGCTTCACACAGCCACCCTTGGGTGATGAGATTTGATCTCAATTAACACCTTTGATGTA		1080
QY	1081	AATACCCCTCCAAACCCAGACCTCTGGCTCCGAGCATGTCCTCCCAAGAAAGCTGTGCTGG		1140
Dp	1081	AATACCCCTCCAAACCCAGACCTCTGGCTCCGAGCATGTCCTCCCAAGAAAGCTGTGCTGG		1140
QY	1141	GAGAGGGCAGAGCTCTCCCATGACAGATGTGTCACAGAAAGGTGCACAAAGTCTCCAAAGCG		1200
Dp	1141	GAGAGGGCAGAGCTCTCCCATGACAGATGTGTCACAGAAAGGTGCACAAAGTCTCCAAAGCG		1200
QY	1201	GAGTCCACGCTCTCCTACTGGGTGGCCACGCCATGTGTGAGAGCTAGATGAGAGAACTGGCT		1260
Dp	1201	GAGTCCACGAGCTCTCCTACTGGGTGGCCACGCCATGTGTGAGAGCTAGATGAGAGAACTGGCT		1260
QY	1261	GGTGCCTTCCTCTCTTCATCCCGGATGGCTGCTGGGACAGAAACAGGCCCAACAGAGACC		1320
Dp	1261	GGTGCCTTCCTCTCTTCATCCCGGATGGCTGCTGGGACAGAAACAGGCCCAACAGAGACC		1320
QY	1321	ACAGCAGAAATCCCGAAGCAGACAGATGTGTGAAGATAGGCTTTGATGGGCTTGGAGGA		1380
Dp	1321	ACAGCAGAAATCCCGTTCACAGATGTGTGAAGATAGGCTTTGATGGGCTTGGAGGA		1380
QY	1381	CGAACAAATTCATCCAGCCTTAGGGACACACCATATCCGACAGCTGCTGTAAATCC		1440
Dp	1381	CGAACAAATTCATCCAGCCTTAGGGACACACCATATCCGACAGCTGCTGTAAATCC		1440
QY	1441	AAGGCCAAGAGTAAACAGAAAGTGAGAAATTAAGACTGTAGTTCTGCCTCCGAGCCCAAG		1500
Dp	1441	AAGGCCAAGAGTAAACAGAAAGTGAGAAATTAAGACTGTAGTTCTGCCTCCGAGCCCAAG		1500
QY	1501	CTGGATTAACCTTCTTATATGTCAGTAAACGGGTGACCAAGTATGTTTGCAATTAATGTGGCCA		1560
Dp	1501	CTGGATTAACCTTCTTATATGTCAGTAAACGGGTGACCAAGTATGTTTGCAATTAATGTGGCCA		1560
QY	1561	AGACCTGGCTAACCGGAAAGTGTTTTGGAAAGATGGCTCTCTTGGAACAGTCCAAAGAAAG		1620
Dp	1561	AGACCTGGCTAACCGGAAAGTGTTTTGGAAAGATGGCTCTCTTGGAACAGTCCAAAGAAAG		1620
QY	1621	ATGCCCCAGAAAACACACTTCTCTGTTCACCTGCGCCCTACACACACTGGGAAGCCACAT		1680
Dp	1621	ATGCCCCAGAAAACACACTTCTCTGTTCACCTGCGCCCTACACACACTGGGAAGCCACAT		1680
QY	1681	GACCAAGTTTACTGTTCCGATCAGACAGGCCCCTAATCCAGTGCAGAGGTTTTGCTTAATAGC		1740
Dp	1681	GACCAAGTTTACTGTTCCGATCAGACAGGCCCCTAATCCAGTGCAGAGGTTTTGCTTAATAGC		1740
QY	1741	TACAACCAAGGTGGCTGAGATCTCTTAATAGAAACAGGGTCAACATGACTCTAA		1800
Dp	1741	TACAACCAAGGTGGCTGAGATCTCTTAATAGAAACAGGGTCAACATGACTCTAA		1800
QY	1801	GTGGATGGGAGTGGTGGAGATCTCTATGCAAGCTGGAAGAACCTTGCGCTTGAATCTCTGC		1860
Dp	1801	GTGGATGGGAGTGGTGGAGATCTCTATGCAAGCTGGAAGAACCTTGCGCTTGAATCTCTGC		1860
QY	1861	CTGCGCTCAGCTTATTTCTTGAATTAATGAGGAGAGGTGTATAGGAAAGTGTGGGGA		1920
Dp	1861	CTGCGCTCAGCTTATTTCTTGAATTAATGAGGAGAGGTGTATAGGAAAGTGTGGGGA		1920
QY	1921	AGTTTTCTGTGTAAATTAAGAGGATCTTTCTTCAA	1958	
Dp	1921	AGTTTTCTGTGTAAATTAAGAGGATCTTTCTTCAA	1958	

RESULT	2
LOCUS	AK012786
DEFINITION	1938 bp mRNA linear HTC 19-JAN-2002
ACCESSION	AK012786
VERSION	AK012786.1 GI:12849758
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (strain:C57BL/6J) 10, 11 days embryo cDNA to mRNA, clone:1b1-Riken full-length enriched mouse cDNA library
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Methods Enzymol. 303, 19-44 (1999)
REFERENCE	99279253
AUTHORS	10349636
TITLE	2 (sites)
JOURNAL	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
REFERENCE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
REFERENCE	20499374
AUTHORS	3 (sites)
TITLE	Shibata, K., Itoh, M., Atzawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitanishi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
JOURNAL	Riken integrated sequence analysis (RISA) system-384-format
REFERENCE	sequencing pipeline with 384 multiplexed sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
REFERENCE	20530913
AUTHORS	11076861
TITLE	4 (sites)
JOURNAL	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
REFERENCE	Functional annotation of a full-length mouse cDNA collection
AUTHORS	Nature 409, 685-690 (2001)
TITLE	5 (bases 1 to 1938)
JOURNAL	Adachi, J., Atzawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Aikawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bull, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagata, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, Y., Kojima, Y., Konno, H., Kouda, M., Koyama, K., Kuthira, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamana, K. I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
JOURNAL	Direct Submission
REFERENCE	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for

|||||
Db 1440 AGAAGTGAATAAAGACGTGTGATTCCTGCCAGACCCAGCCAGATGATACCTTTAT 1499
QY 1517 GTGAGTGAAGGTCAGCAGAGTGTATTTGCAATTAGTGGGCCAAGACCTGCGTAACCG 1576
Db 1500 GTGAGTGAAGGTCAGCAGAGTGTATTTGCAATTAGTGGGCCAAGACCTGCGTAACCG 1559
QY 1577 AGTGTATTTGGAAGATGAGTCTCTTGAGCAGATCCAGAGAGATGCCAGAAACACA 1636
Db 1560 AAGTGTATTTGGAAGATGAGTCTCTTGAGCAGATCCAGAGAGATGCCAGAAACACA 1619
QY 1637 CTTCCTGTGTACACTGCGCCCTGCACACACTGGGAAGCCACATGACAGTTACTGTTC 1696
Db 1620 CTTCCTGTGTACACTGCGCCCTGCACACACTGGGAAGCCACATGACAGTTACTGTTC 1679
QY 1697 CGATCAGCAGGGCTACTCTCAGTTCAGAGGTTTGTATAGCTAACACAGAGTGGCC 1756
Db 1680 CGATCAGCAGGGCTACTCTCAGTTCAGAGGTTTGTATAGCTAACACAGAGTGGCC 1739
QY 1757 TGCAGCTCTTTTGTATTAAGAACAGGGTACATTAAGTAAAGGATGGAGTCTG 1816
Db 1740 TGCAGCTCTTTTGTATTAAGAACAGGGTACATTAAGTAAAGGATGGAGTCTG 1799
QY 1817 GAGATCTTATGACAGCTGAGAGACCTGCGCTTGAACCTCGCTGCTCCAGCTTAT 1876
Db 1800 GAGATCTTATGACAGCTGAGAGACCTGCGCTTGAACCTCGCTGCTCCAGCTTAT 1859
QY 1877 GCTTGAATTAATGAGGAGAGTGTGATTAAGGAAAGTTGGGAAAGTTTCTGTATAA 1936
Db 1860 GCTTGAATTAATGAGGAGAGTGTGATTAAGGAAAGTTTCTGTATAA 1919
QY 1937 TAAAAAGGATCTTTCTT 1955
Db 1920 TAAAAAGGATCTTTCTT 1938

RESULT 3
AL560912 887 bp mRNA linear EST 16-FEB-2001
LOCUS AL560912 LTI_NFL010_BC2 Homo sapiens cDNA clone CS0DL005Y108 5
DEFINITION prime, mRNA sequence.
ACCESSION AL560912
VERSION AL560912.1 GI:12907828
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 887)
AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 101 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. 887
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DL005Y108"
/clone_1id="LTI_NFL010_BC2"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/note="Vector: pCMVSPORT 6, Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Full prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9600 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :

BASE COUNT 251 a 219 c 245 g 171 t 1 others
ORIGIN

Query Match 34.0%; Score 671.2; DB 9; Length 887;
Best Local Similarity 87.2%; Pred. No. 1,4e-128;
Matches 769; Conservative 1; Mismatches 109; Indels 3; Gaps 3;

|||||
QY 96 CTTCAGTCGAGCCATCATGCTATCTCTCTCTGTGACATATGCTCCAGACTTCTTCA 155
Db 4 CTTCAGTCGAGCCATCATGCTATCTCTCTGTGACATATGCTCCAGACTTCTTCA 63
QY 156 TCACTCCCGTACGCTGAGCTCCATCCACTGTGGCCACACTTTTCATGCAATGCTAAT 215
Db 64 TCACTCCCGGAGCTGGCCCGCCATCCACTGTGGCCACACTTTTCATGCAATGCTAAT 123
QY 216 CCACTGTTTGAAGACAGCACCACCAAGTGGACCTGCCACAGTGAATCCAGTTGGCAA 275
Db 124 TCACTGTTTGAAGACAGCACCACCAAGTGGACCTGCCACAGTGAATCCAGTTGGCAA 183
QY 276 AAGACTATTATTAACAACCTTTCTTGACCTTGCCCGAAGAGAGATGCTTGA 335
Db 184 AAGACTATTATTAACAACCTTTCTTGACCTTGCCCGAAGAGAGATGCTTGA 243
QY 336 TGCAGATTTCTTAAGATGAAGTGAAGAGTCAAGCTCAGCTTCCAGAAAGACAG 395
Db 244 TGCAGATTTCTTAAGATGAAGTGAAGAGTCAAGCTCAGCTTCCAGAAAGACAA 303
QY 396 GGAGAAACGGGACAGCCAGGCCATTATGACACTCTACGGGACACCTGGAAGAACCAA 455
Db 304 GGAGAAACGGGACAGCCAGGCCATTATGACACTCTACGGGACACCTGGAAGAACCAA 363
QY 456 TGTACCGTGTAGTCCCTACAGAACCGCTTAACAAGCAGAGTCTGTCTTCCACCT 515
Db 364 TGTACCGTGTAGTCCCTACAGAACCGCTTAACAAGCAGAGTCTGTCTTCCACCT 423
QY 516 GAAAGAAACAGATGAAGTCTCTGAGCAGCGCAGAGTGAAGACCAAGCTCGGAGGA 575
Db 424 GAAAGAAACAGATGAAGTCTCTGAGCAGCGCAGAGTGAAGACCAAGCAGAGGA 483
QY 576 GGCCACCAGCTCAAGTGAAGATGAAGAAACCATGAGCAATGAGTCTTACTCCAGAG 635
Db 484 GGCCACCAGCTCAAGTGAAGATGAAGAAACCATGAGCAATGAGTCTTACTCCAGAG 543
QY 636 CCAGCGTTCTGAGGTGAGAGATGATTCAGAGATGGGTGGGACAGTACAGGCGTGA 695
Db 544 CCAGCGCGCTGAGGTGAGAGATGATTCAGAGATGGGTGGGACAGTACAGGCGTGA 603
QY 696 GCAGCTGGCTGTACTGCGCTCCCTCAAGAAAGATGAGATGCAAGGAGCTCG 755
Db 604 ACAGCTGGCTGTACTGCTGTGTCTCTCAAGAAAGATGAGAGATCTTAAAGAGGACAG 663
QY 756 GAAGGCCACAGAGGAGTGTGAGAGTGAAGAGTTGGTGTCTTGAAGACA 814
Db 664 GAAGGCCCTGAGGAGGAGTGTGAGAGTGAAGAGTTGGTGTCTTCCAGAGACA 723
QY 815 AGTTGAAGACTCTCAACACTGAGCTGATGAGCGGCAACTTGAAGTCAAGTCAAGCCAGA 874
Db 724 AGTTGAAGACTCTCAACACTGAGCTGATGAGCGGCAACTTGAAGTCAAGTCAAGCCAGA 783
QY 875 AGGACTTAACAAGTGTGAGCAGAGATGAGCAGCTTAAGAAAGTCTGATGATCTC 934
Db 784 AGGACTTAACAAGTGTGAGCAGAGATGAGCAGCTTAAGAAAGTCTGATGATCTC 842
QY 935 CAGGGAACCTTGAAGCTGCTCC-GCAGCAATGAGACGCTC 975
Db 843 CAGGGAACCTTGAAGCTGCTCCAGCAGTGGSCAGTGAAGTCTC 884
RESULT 4
AL560947 959 bp mRNA linear EST 16-FEB-2001
LOCUS AL560947

[illegible]

Db	477	GCTGTGCTCCACACTGAAAAAGCAGATGAAGTAACTTAGAGCGACGACGACGATGACCCA	536
Oy	561	ACAAGCTCGGAGGAGGAGGCCACCAGCTCAAGTCAGATGAATAAACCATGGACAAATTGA	620
Db	537	ACAGAACAAGAGGAGGAGGCCCGGCTCAGGAGCAAGATGAAGACCATGGAGCAGATTGA	596
Oy	621	GCTCCTACTCCAGAGGCACGCTTTGTAGTGAGTGAGAGATGATTCAGACATGGGTGGG	680
Db	597	GCTTCTACTCCAGAGGCACGCTTTGTAGTGAGTGAGAGATGATTCAGACATGGGTGGG	656
Oy	681	ACAGTCAGGAGGAGGACGACCTGGCTGTGTACGTGCGGTCCGTAAGAAAGATGAGAA	740
Db	657	ACAGTCAGGAGGAGGACGACCTGGCTGTGTGTGTGTCTTCAGAAAAGATGAGAGAA	716
Oy	741	TCTGAAGAAGCTCGGAAGGCCACAGAGGAGAACCTGGCTGACAGGTTGAAAGAGATTGGT	800
Db	717	TCTAAGAAGGAGGACGAGAACGCTTGAAGGAGAGGTGGCTGACA-RCGTGAGGAAGATTGT	775
Oy	801	GTCCTCTAGGAGCAAGTTGAGACTCTCAACACTGAGCTGATTCAGAGCCAGTTAGACT	860
Db	776	TTCTCTCCAGAACCAAGTTTCAGACAGTCTACCTGATGTGATGATCAGGCAAGTTAGAACT	835
Oy	861	GAGGTCAGCCCCAGGAAGGACTTCAAAGTGTCTTACACAGAGATTCAGAGGCTTAAGAAAGA	920
Db	836	GAGTCAGCCCCAGGAAGGACTTCAAAGTGTCTTAAAGGAATCATGAGCTTGAAGAAAGA	895
Oy	921	GCTCATGATGCTCCAGGAGAACCTTGAGCGCTGCTTC-CGCGACCAATGAGAGCGTCAGCC	979
Db	896	GCTAACGATGCTGCAAGAAACCTTGAGAACCTGCCACAAAGTGGSCAGTAGACTGTGCAGCC	955
Oy	980	GCCT 983	
Db	956	GCCT 959	

RESULT 5
B1151643

LOCUS	B1151643	702 bp	mRNA	linear	EST 05-JUL-2001
DEFINITION	602916189P1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:5066423 5', mRNA sequence.				
ACCESSION	B1151643				
VERSION	B1151643.1 GI:14611644				
KEYWORDS	EST.				
SOURCE	house mouse. Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 702) NIH-MGC htp://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC)				
AUTHORS	Unpublished (1999)				
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgapbs@email.nih.gov				
COMMENT	Tissue Procurement: Gilbert Smith, Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: LLM1179 row: b column: 24 High quality sequence stop: 694.				

FEATURES

source location/Qualifiers

1..702

/organism="Mus musculus"

/strain="CZECH II (fetal)"

/db_xref="taxon:10090"

/clone="IMAGE:5066423"

/clone_lib="NCI CGAP Lu29"

/issue_type="spontaneous tumor, metastatic to mammary stem cell origin."

/note="Organ: Lung; Vector: pCMV-SPORT6; Site:1; SalI; /lab_host="DH10B"


```
QY 486 GAGATGCTGTGTTCCACCTTGAAAAACAGATGAAGTTCCTGGACGCGCAGATGAG 555
|||||
DB 481 GAGATGCTGTGTTCCACCTGTAATAACAGATGAAGTTCCTGGACGCGCAGATGAG 540
|||||
QY 556 ACCAACAAGCTCGGAGAGGCGCCACGACTCAAGTCAGATGAAAAACCATGAGCAA 615
|||||
DB 541 ACCAACAAGCTCGGAGAGGCGCCACGACTCAAGTCAGATGAAAAACCATGAGCAA 600
|||||
QY 616 ATTGAGCTCTCTACCTCAGAGCGGCTCTGAGTGGAGGAGATGATTCAGACATGGGT 675
|||||
DB 601 ATTGAGCTCTCTCTACCTCAGAGCGGCTCTGAGTGGAGGAGATGATTCAGACATGGGT 660
|||||
QY 676 GTGGACAGTACGCGGTGAGCAGCTG 702
|||||
DB 661 GTTGACAGTACGCGGTGAGCAGCTG 687
|||||

RESULT 7
BF162255 744 bp mRNA linear EST 30-OCT-2000
LOCUS 60170948P1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3990371 5',
DEFINITION mRNA sequence.
ACCESSION BF162255
VERSION BF162255.1 GI:11042462
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 744)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: L1AM9201 row: 0 column: 12
High quality sequence stop: 730.
Location/Qualifiers
1. 744
/organism="Mus musculus"
/strain="C2EHC II (fetal)"
/db_xref="taxon:10090"
/clone="IMAGE:3990371"
/clone_lid="NCI CGAP Lu29"
/tissue_type="spontaneous tumor, metastatic to mammary.
Stem cell origin."
/lab_host="DHI0B"
/note="Organ: Lung; Vector: pCMV-SPORT6; site_1: SalI;
site_2: NotI; Cloned unidirectionally. Primer: oligo dT.
Library constructed by: Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 201 a 183 c 212 g 148 t
ORIGIN
```

```
QY 141 CTCGACTTCTTCGATCATCCTCCGAGACGTGGTCCCATCCACTGTGGCCACTTTTCA 200
|||||
DB 121 CTCGACTTCTTCGATCATCCTCCGAGACGTGGTCCCATCCACTGTGGCCACTTTTCA 180
|||||
QY 201 TCTGCATGCTTAATCCAGTGTGTGAGACAGCACCAAGTGGACCTGCCACAGTGTAG 260
|||||
DB 181 TCTGCATGCTTAATCCAGTGTGTGAGACAGCACCAAGTGGACCTGCCACAGTGTAG 240
|||||
QY 261 AATCCAGTGTGGCAAAAAGACTATTATTAACAACCTTTTCTTACCTGCCAGAGA 320
|||||
DB 241 AATCCAGTGTGGCAAAAAGACTATTATTAACAACCTTTTCTTACCTGCCAGAGA 300
|||||
QY 321 GGAGATGCTTGGATGAGATTCCTTAAGAATGAAGTGGACAGCGTCAAGCTCAGCT 380
|||||
DB 301 GGAGATGCTTGGATGAGATTCCTTAAGAATGAAGTGGACAGCGTCAAGCTCAGCT 360
|||||
QY 381 TTCCAGAAAGACAGGAGAAACGGACAGCCAGCCATTATCGACACTACGGGACAC 440
|||||
DB 361 TTCCAGAAAGACAGGAGAAACGGACAGCCAGCCATTATCGACACTACGGGACAC 420
|||||
QY 441 CCTGGAGAACGCAATGCTACCTGGAGTCCCTACAGACGCTTTAAACAAGCAGAGAT 500
|||||
DB 421 CCTGGAGAACGCAATGCTACCTGGAGTCCCTACAGACGCTTTAAACAAGCAGAGAT 480
|||||
QY 501 GCTGTGTTCCACCCGAAAAACAGATGAAGTTCCTGGAGCAGCGGAGATGAGACCA 560
|||||
DB 481 GCTGTGTTCCACCCGAAAAACAGATGAAGTTCCTGGAGCAGCGGAGATGAGACCA 540
|||||
QY 561 ACAAGCTCGGAGAGAGGCCACCGACTCAAGTGCAGATGAAACCATGAGACAAATTGA 620
|||||
DB 541 ACAAGCTCGGAGAGAGGCCACCGACTCAAGTGCAGATGAAACCATGAGACAAATTGA 598
|||||
QY 621 GCTCTACTCTCAGAGCCAGCGTTCTGAGGTGGAGAGATGTTGAGACATGGGTGGG 680
|||||
DB 599 GCTCTACTCTCAGAGCCAGCGTTCTGAGGTGGAGAGATGTTGAGACATGGGTGGG 658
|||||
QY 681 ACAGTCAGCGGTGAGAGCGTGGCTGTACTGCTGCTCCCTCAAGAAAGATGATGAA 740
|||||
DB 659 ACAGTCAGCGGTGAGAGCGTGGCTGTACTGCTGCTCCCTCAAGAAAGATGATGAA 715
|||||
QY 741 TCTGAGAGA 749
|||||
DB 716 TTGAAGAA 724
|||||

RESULT 8
BM464844 1061 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6428661 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5504946
DEFINITION 5', mRNA sequence.
ACCESSION BM464844
VERSION BM464844.1 GI:18513886
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1061)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: L1AM12147 row: b column: 19
High quality sequence start: 2
```


Db 240 CCAGCCCAAGCTGATACCTTCCTTATGTCACTGACGAGGTGACAGAGTCATGTTGCAAT 299
OY 1551 TAGTGGCCCAAGACCTGGCTTAACCCGGAAGTGTGTTGGAAGATGGCTCCTCTTGGACAG 1610
Db 300 TAGTGGCCCAAGACCTGGCTTAACCCGGAAGTGTGTTGGAAGATGGCTCCTCTTGGACAG 358
OY 1611 TCAGAGAGATGCCCCAGAAAAACACACATTCCTGTTGACAGCCGCTGACACACACTGG 1670
Db 359 TCAGAGAGATGCCCCAGAAAAACACACATTCCTGTTGACAGCCGCTGACACACACTGG 418
OY 1671 G-AAGCCCAATGACCACTTACTGTTCCGATCAGACAGGCGCTACTCCAGTTGCAAGGTT 1729
Db 419 GAAACCCACATGACCACTTACTGTTCCGATCAGACAGGCGCTACTCCAGTTGCAAGGTT 478
OY 1720 TTGCTTAAGCTACCAACAGGTGTGCTGACCTCTTTGTTTATAGAACAGAGGTAC 1789
Db 479 TTGCTTAAGCTACCAACAGGTGTGCTGACCTCTTTGTTTATAGAACAGAGGTAC 538
OY 1790 ATTGACTTAAGTGGATGGAGTGTGAGATCCTATGCAAGGCTGAGAGACCTGCGCT 1849
Db 539 ATTGACTTAAGTGGATGGAGTGTGAGATCCTATGCAAGGCTGAGAGACCTGCGCT 598
OY 1850 TGAACCTCTG 1859
Db 599 TGAACCTCTG 608

RESULT 10
LOCUS BE334637 589 bp mRNA linear EST 14-JUL-2000
DEFINITION us01912.y1 NCI-CGAP_Mam4 Mus musculus cDNA clone IMAGE:3257638 5' sequence.
ACCESSION BE334637
VERSION BE334637
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 589)
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Unpublished (1997)
TUMOR INDEX
CONTACT: Robert Strausberg, Ph.D.
JOURNAL Email: cgaps-remail.nih.gov
COMMENT Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furch Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: lmage.lnl.gov/image/html/lresources.shtml
MGI:1069002
Seq primer: -40RP from Gibco
High quality sequence stop: 362.
Location/Qualifiers
1. 589
FEATURES
Source
/organism="Mus musculus"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:3257638"
/clone_id="NCI-CGAP_Mam4"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Salt; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furch,

Query Match 27.3%; Score 539; DB 10; Length 589;
Best Local Similarity 97.8%; Pred. No. 3e-101;
Matches 578; Conservative 0; Mismatches 10; Indels 3; Gaps 3;
ORIGIN
BASE COUNT 152 a 151 c 172 g 114 t
NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996).
Query 681 ACAGTCACGGGTGAGACAGCTGCTGTACTGCTGCTCCCTCAAGAAAGTATGAGAA 740
Db 1 ACAGTCACGGGTGAGACAGCTGCTGTACTGCTGCTCCCTCAAGAAAGTATGAGAA 60
OY 741 TCTGAGAGACCTGGAGAGGCACAGGGAGCTGGCTGACAGGTTGAAGAGATTGGT 800
Db 61 TCTGAGAGACCTGGAGAGGCACAGGGAGCTGGCTGACAGGTTGAAGAGATTGGT 120
OY 801 GTCCCTAGAGACCAAGTTGAAGACTCTACACTGAGCTGATCAGGCCAAGTTAGACT 860
Db 121 GTCCCTAGAGACCAAGTTGAAGACTCTACACTGAGCTGATCAGGCCAAGTTAGACT 180
OY 861 GAGTCAGCCCGAAGACCTTACAAAGTCTGACACAGAGATCAAGACCTTAAGAAAG 920
Db 181 GAGTCAGCCCGAAGACCTTACAAAGTCTGACACAGAGATCAAGACCTTAAGAAAG 240
OY 921 GTCGATGATCTCCAGGAGACCTTGAAGCCCTGCC-CCGACCAATGAGAGGTACGCC 979
Db 241 GTCGATGATCTCCAGGAGACCTTGAAGCCCTGCCGACCAATGAGAGGTACGCC 299
OY 980 GCCTGCTTTTGAAGAGCCACCCCTGTGAGATGATGAACCCGAGCTTACACAGCCAC 1039
Db 300 GCCTGCTTTTGAAGAGCCACCCCTGTGAGATGATGAACCCGAGCTTACACAGCCAC 359
OY 1040 CCTTCGGTATGATGATGATTCATATCACACCTTTGATTAATACCCCTCAACCCAGA 1099
Db 360 CCTTCGGTATGATGATGATTCATATCACACCTTTGATTAATACCCCTCAACCCAGA 419
OY 1100 CCTTCGGTATGATGATGATTCATATCACACCTTTGATTAATACCCCTCAACCCAGA 1159
Db 420 CCTTCGGTATGATGATGATTCATATCACACCTTTGATTAATACCCCTCAACCCAGA 479
OY 1160 TGCAGAAATGCTCCAGAGAGGTGACCAAAAGTCTCCAAAGCCGAGTCCAGCTCTAC 1219
Db 480 TGCAGAAATGCTCCAGAGAGGTGACCAAAAGTCTCCAAAGCCGAGTCCAGCTCTAC 539
OY 1220 GTGGCCAGCAATGCTGTAGAGAGCTAGATGAGAACTGCTGCTTC 1270
Db 540 GTGGCCAGCAATGCTGTAGAGAGCTAGATGAGAACTGCTGCTTC 589

RESULT 11
LOCUS AA684194 554 bp mRNA linear EST 09-DEC-1997
DEFINITION vm08410.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone IMAGE:1003411 5' similar to TR:008854 008854 MTRIP. ;, mRNA sequence.
ACCESSION AA684194
VERSION AA684194
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 554)
REFERENCE Maria, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Maria M/Mouse EST Project

WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: mousestewartson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:567627

Possible reversed clone: similarity on wrong strand
High quality sequence stop: 512.

FEATURES

source

1..554
Location/Qualifiers
/organism="Mus musculus"
/strain="B6D2 F1/J"
/db_xref="taxon:10090"
/clone_id="Knowles Solter mouse 2 cell"
/tissue_type="embryo"
/dev_stage="2-cell"
/lab_host="DH10B"
/note="Organ: embryo; Vector: pBluescribe (modified);
Site_1: MluI; Site_2: SalI; Cloned unidirectionally from
mRNA prepared from 13,500 2-cell stage embryos. Primer:
SalI(dT): 5'-CGTCGACCGTCGACCGCTTTT-3'. CDNA
were cloned into the MluI/SalI sites of a modified
pBluescribe vector using commercial linkers (NEB).
Average insert size: 1.2 kb."

BASE COUNT
ORIGIN

Query Match 26.9%; Score 532.2; DB 9; Length 554;

Best Local Similarity 99.3%; Pred. No. 7.8e-100;
Matches 545; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 53 CCACCAACTGTCTGTCTGTGCGAGCTGGTCCCTGGGCTGCTTGAAGTGCAGCATCA 112
DB 6 CCACCAACTGTCTGTCTGTGCGAGCTGGTCCCTGGGCTGCTTGAAGTGCAGCATCA 65
QY 113 TGCCTATCCTCTCTGTGTGACATCTGTCTCCGACTTTTGATCATTCCCGTGCATGG 172
DB 66 TGCCTATCCTCTCTGTGTGACATCTGTCTCCGACTTTTGATCATTCCCGTGCATGG 125
QY 173 CTGGCATCCACTGTGGGCGACACTTTTCATCTGCAATGCTAATCCAGTGGTTTGACAG 232
DB 126 CTGGCATCCACTGTGGGCGACACTTTTCATCTGCAATGCTAATCCAGTGGTTTGACAG 185
QY 233 CACCAAGTCGAGCTGCCACAGTGTAGATCCAGGTGGGCAAAAGACTATTATAACA 292
DB 186 CACCAAGTCGAGCTGCCACAGTGTAGATCCAGGTGGGCAAAAGACTATTATAACA 245
QY 293 AACCTTTTCTTGACCTGCCCGACGAGAGAGATGTTTGGATGACAGATTTCTTAAGA 352
DB 246 AACCTTTTCTTGACCTGCCCGACGAGAGAGATGTTTGGATGACAGATTTCTTAAGA 305
QY 353 ATGAAGTCGAGCTGCAAGCTGCTTCCAGAAAGACAGGAGAAAGGCGACAGCC 412
DB 306 ATGAAGTCGAGCTGCAAGCTGCTTCCAGAAAGACAGGAGAAAGGCGACAGCC 365
QY 413 AGGCATTATGACACTCTAGGGGACACCTGGAAGAACGCAATGCTACCTGTGAGTCC 472
DB 366 AGGCATTATGACACTCTAGGGGACACCTGGAAGAACGCAATGCTACCTGTGAGTCC 425
QY 473 TACGAAGAGCTTAAACAAGCAGAGATGCTGTTCACACCTT-GAAAAACAGATGAAG 531
DB 426 TACGAAGAGCTTAAACAAGCAGAGATGCTGTTCACACCTTGAAGAAAAACAGATGAAG 485
QY 532 TTCTGTGAGCAGCGCAGAGATGAGACCAAAAGCTGGGAGGAGGCCACCGACTCAAG 591
DB 486 TTCTGTGAGCAGCGCAGAGATGAGACCAAAAGCTGGGAGGAGGCCACCGACTCAAG 545
QY 592 TGCAAGATG 600

DB 546 TGCAAGATG 554

RESULT 12

LOCUS

BG764118 702 bp mRNA linear EST 15-MAY-2001
60273137F1 NIH_MGC_49 Homo sapiens CDNA clone IMAGE:4862312 5',

DEFINITION

mRNA sequence.

ACCESSION

BG764118 GI:14074771

VERSION

EST.

KEYWORDS

human.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCU/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LCM1722 row: 3 column: 09
High quality sequence stop: 697.

FEATURES

source

1..702
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="IMAGE:4862312"
/clone_id="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using Zap-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

BASE COUNT
ORIGIN

Query Match 26.0%; Score 514.4; DB 10; Length 702;

Best Local Similarity 87.4%; Pred. No. 3.6e-96;
Matches 574; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

QY 81 TGGTCCCTGGGCTGCTGAGTCGAGCATCATGCTTCTCTCTGTGACATATG 140
DB 42 TGGTCCCTGGGCTGCTGAGTCGAGCATCATGCTTCTCTCTGTGACATATG 101
QY 141 CTCCGACTTCTTGCATCTCCCGTGCAGTGGTCCATCCACTGTGGCCACACTTTTCA 200
DB 102 CTCCGACTTCTTGCATCTCCCGTGCAGTGGTCCATCCACTGTGGCCACACTTTTCA 161
QY 201 TCTGCAATGCTTAATCCAGTGTGTTGAGACAGCACCAGAGTGGAGCTGCCACAGTGA 260
DB 162 CTGCAAGTGCCTTAATTCAGTGTGTTGAGACAGCACCAGAGTGGAGCTGCCACAGTGA 221
QY 261 AATCAGGTTGGCAAAAGACATTAATTAACAATTTTCTTGGACCTCGGCCAGGAGA 320
DB 222 AATCAGGTTGGCAAAAGACATTAATTAACAATTTTCTTGGACCTCGGCCAGGAGA 281
QY 321 GGAGAAATGCTTGGATGCGAATTTCTTAAGAATGAAGTGCAGAGCTCAAGCTCAGCT 380

Db 282 GGAGGATATGCTTGATGATGCAGATTCTTTAAAGATATACGTGGACCAATGTGAGAGCCAGCT 341

QY 381 TTCCGAGAAAGACAGGGAGAAAAGGGAGACCCAGGCCATTATCGACACTCTACGGGACAC 440

Db 342 TTCCGAGAAAGACAAAGGAGAAAACGAGACAGCCAGGTTCATCATCGACACACTCTCGGGATAC 401

QY 441 CCTGGAGAGCCCAATGTGTACCGTGGAGTCCCTACAGAAAGCCCTTAACCAAGGCAAGAT 500

Db 402 GCTGGAAAGACCAATGTCTACTGTGTGTATCTTCGACGACGCGCTTGGGAGAGCCGAGAT 461

QY 501 GCTGTGTTCCACCCCTGAAAAAAACAGATGAAGTTCTCTGGAGCAGCGGCAAGATGAGACCAA 560

Db 462 GCTGTGCTCCACACTGAAAAAAGCAATGAAAGTACTTAGAGCAGCAGCAAGGATGAGACCAA 521

QY 561 ACAACCTGGGGAGGGAGGCCACCGACTCAAGTCGAAGTGAAGAAACCATTGAGACCAATTGA 620

Db 522 ACAACCAACAGGGAGGGCCCGCGGCTCAGGAGCAAGATGAAGACCAATGACAACATTGA 581

QY 621 GCTCTACTCCAGAGACCAGCGTTCTTGAGGTGAGAGA-GATGATTCGAGACATGAGTGTGG 679

Db 582 GCTTCTACTCCAGAGACCAGCGCCCTTGAGGTGAGAGCATGATCCGAGACATGGGTGTGG 641

QY 680 GACACTCAGCGGTGAGAGACGTGCGTGTGTACTGCGTGTCTCAAGAAAGACTG 736

Db 642 GACACACAGGGGTGGACACAGCTGGGTGTGTGTAAGTGTGTCTCAAGAAAGGTACG 698

RESULT	13
LOCUS	BC682548
DEFINITION	BC682548 752 bp mRNA linear EST 01-MAY-2001 60268236221 NCI-CGAP-Skn4 Homo sapiens CDNA clone IMAGE:4749447 5'
ACCESSION	mRNA sequence.
VERSION	BC682548
KEYWORDS	BC682548.1 GI:13913945
SOURCE	EST. human.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 752)
NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

cDNA library preparation: Life Technologies, Inc.
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLNL0602 row: k column: 16
 High quality sequence start: 19
 High quality sequence stop: 736.

FEATURES

source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4749447"
/clone_1lb="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: Skin; Vector: pCW-SPORT6; Site:2: Salt; Cloned unidirectionally. Primer: Oligo dT Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP Library."

```

Query Match

26.08; Score 513.6; DB 10; Length 752;

Best Local Similarity 84.5%; Pred. No. 5, 2e-96;
Matches 612; Conservative 0; Mismatches 109; Indels 3; Gaps 3;

OY	52	TCACCAAACTGTGTGTCTCTGTGGACAGTGGTCCCTGGGCTGCTTGAATCCAGGCATC	1111
Db	27	TCACCAAGCCGGACACTGTACACTTTCTTTTGGCTGCTGGGGCCCTTGTAGTCCAGCTATC	86
OY	112	ATGCCTATCTCTCTCTGTGTGACTATGTGCTCCGACTTCTTTCGATCTCCCGGACGTG	1717
Db	87	ATGCCTATACGTGCTGTGTGACTATGTGCTCCGACTTCTTTCGATCTCTCCGGACGTG	1466
OY	172	GCTGCCATCCACTGTGGCCACACTTTTTCATCTGCAATGCCATTCAGTGGTTTGAGACA	2311
Db	147	GCCGCAATCCACTGTGGCCACACTTTCCTCCACTTGGCAGTGCCATATTCAGTGGTTGAAACA	2066
OY	232	GCACCAATGTGGACCTGGCCCACTGTGTAAGATCCAGG-TTGGCAAAAAGACTTTATAAA	2900
Db	207	GCACCAATGTGGACCTGGCCCACTGTGTAAGATCCAGGTTTGGCAAAAAGACTTTATCAA	2666
OY	291	CAAACTTTTCTTTTGAACCTCGCCACAGAAAGAGAAATGTCTTGATGCAGAAATTCCTTAAA	3500
Db	267	TAAGCTCTTCTTTGATGCTGTGGCCACAGAGAGAGAAATGTCTTGGATGCAGAAATTCCTAAA	3266
OY	351	GAATGAACTGACAGCGTCAAAGCTCAGCTTCCACAGAAAGACAGGAGAAACGGGACAG	4100
Db	327	GAATGAACTGACAGATGTCAAGGCCACGCTTCCCAAGAAAGACAGGAGAAACGGGACAG	3866
OY	411	CCAGGCCATTTATCCACACTCTACGGGGACACCTCTGGAAGAACGCATCTACCTGGAGTGC	4700
Db	387	CCAGGCTCTATCTATCCAGACACTCTGCGGGATAGCCTGTGAAGAACGCATCTCTACTGTGGTATC	4466
OY	471	CCCTACAGAAACCCTTTAAACAAGAGAGAGATGCTGTGTTCCACCCCTGAAAAAACAAGATGAA	5300
Db	447	TCTGACAGACGCTTTGGGCAAGGCCAGATGTCTGTCTCCACACTGAAAAAACAAGATGAA	5066
OY	531	GTTCCCTGGAGCAGCGGCAGATGAGACCAACAAGCTCGGGAGAGGGCCCAACGCACTCAA	5900
Db	507	GTACTTAGAGCAGCAGCAGATGAGACCAACAAGCACAAGAGAGAGGCCCGCGGCTCAG	5666
OY	591	GTCGAAGATGAAAAACATGTGAGGCAAAATTTGAGTCTCTACTCCAGAGCCAGCGTTCTGAGT	6500
Db	567	GAGCAAGATGGAAGCCATGTGAGGCAAGATTTTGAATCTCTCTCCAGAGCCAGCGCCCTGAGGT	6266
OY	651	GGAGAGATGATTTGAGACATAGGTGTGGGACACTCAGCGGTGG-AGCAGCTGGCTGTGT	7090
Db	627	GGAGAGATGATTTCCGAGACATAGGTGTGTGGGACACAGCGGTGGAAACAAGTGGCTGTGT	6866
OY	710	ACTGCGTGTCCCT-CAAGAAAAGATATGAAATCTGGAAGAAAGCTCGGAAGGCCACAGGG	7688
Db	687	ACTGTGTGTCTCTCCAGAAAAGATAGCAGAAATCTTAAAAAAGAGCGCGGAAGGCTCCGGG	7466
OY	769	GAAC 772	
Db	747	GCAC 750	

RESULT
BE30418

LOCUS	BE304189	519 bp	MRNA	linear	EST 13-JUL-2000
DEFINITION	6010866421F1 NCT_CGAP_Mam6 Mus musculus CDNA clone IMAGE:3500621 5',				
mRNA sequence.					

ACCESSION	BE304189
VERSION	BE304189.1
	GI:9174333

KEYWORDS	EST.
SOURCE	house mouse.

ORGANISM Mus musculus
Eukaryota; M

REFERENCE
1 (pages 1 to 519)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health. Mammalian Gene Collection (MGC)

JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.

CONTACT: ROBERT STANISBERG, PH.D.

Email: c9abps-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LHM8559 row: 1 column: 06
High quality sequence stop: 515.

FEATURES

source

Location/Qualifiers

1. 519

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone_image="3500621"

/clone_lib="NCI_CGAP_Mam6"

/sex="Female, virgin"

/tissue_type="infiltrating ductal carcinoma"

/dev_stage="5 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT

159 a 116 c 152 g 92 t

ORIGIN

Query Match

26.0%; Score 513.4; DB 10; Length 519;

Best Local Similarity 99.8%; Pred. No. 6e-96;

Matches 514; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 263 TCCAGGTTGGCAAAAAGCTATTATTAACAACCTTTCTTGACCTGCCCGAGAGAGG 322

Db 5 TCCAGGTTGGCAAAAAGCTATTATTAACAACCTTTCTTGACCTGCCCGAGAGAGG 64

QY 323 AGAATGCTTGGATGACAGATTCTTAAAGATGAACGTGACAGCGTCAAGCTCAGCTTT 382

Db 65 AGAATGCTTGGATGACAGATTCTTAAAGATGAACGTGACAGCGTCAAGCTCAGCTTT 124

QY 383 CCCAGAAAGACAGAGAGAAACGGGACAGCCAGCCATTATGACACTCTACGGGACACC 442

Db 125 CCCAGAAAGACAGAGAGAAACGGGACAGCCAGCCATTATGACACTCTACGGGACACC 184

QY 443 TGGAGAAACGCAATGCTACCTGGAGTCCCTACAGAACGCTTAAACAAGCAGAGATGC 502

Db 185 TGGAGAAACGCAATGCTACCTGGAGTCCCTACAGAACGCTTAAACAAGCAGAGATGC 244

QY 503 TGTGTTCCACCTGAAAAAAGATGAAGTCTCGAGACAGCGGAGATGAGACCAAG 562

Db 245 TGTGTTCCACCTGAAAAAAGATGAAGTCTCGAGACAGCGGAGATGAGACCAAG 304

QY 563 AAGCTCGGAGAGAGCCACGACTCAAGTCAAGATGAAAAACATGAGCAAAATTGAGC 622

Db 305 AAGCTCGGAGAGAGCCACGACTCAAGTCAAGATGAAAAACATGAGCAAAATTGAGC 364

QY 623 TCCCTACTCAAGCCAGGTTCTGAGGTGAGAGAGATATTCGAGACATGGGTGGGAC 682

Db 365 TCCCTACTCAAGCCAGGTTCTGAGGTGAGAGAGATATTCGAGACATGGGTGGGAC 424

QY 683 AGTAGCGGTGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 742

Db 425 AGTAGCGGTGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 484

QY 743 TGAAGGAAGCTCGAAGGCCACAGGGAACTGGCT 777

Db 485 TGAAGGAAGCTCGAAGGCCACAGGGAACTGGCT 519

RESULT 15

Bg120736

Bg120736

784 bp mRNA linear EST 30-JAN-2001

LOCUS

DEFINITION 602347389F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4442038 5',
mRNA sequence.
ACCESSION Bg120736
VERSION Bg120736.1 GI:12614245
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Email: c9abps-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LHM10214 row: b column: 23
High quality sequence stop: 713.

Location/Qualifiers

1. 784

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="4442038"

/clone_lib="NIH_MGC_90"

/tissue_type="adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: Liver; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; Oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT

207 a 200 c 224 g 153 t

ORIGIN

Query Match

25.9%; Score 510.6; DB 10; Length 784;

Best Local Similarity 87.2%; Pred. No. 2.2e-95;

Matches 594; Conservative 0; Mismatches 84; Indels 3; Gaps 3;

QY 81 TGTTCCTGGGCTGCTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 140

Db 62 TGTTCCTGGGCTGCTTGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 121

QY 141 CTGCGACTTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 200

Db 122 CTGCGACTTCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 181

QY 201 TCTGCAATGCTTAATCAAGTGTGAGACAGACCAAGTGGAGTGGAGTGGAGTGGAGTGGAG 260

Db 182 CTGCAATGCTTAATCAAGTGTGAGACAGACCAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 241

QY 261 AATCCAGTGTGGCAAAAAGACTATTATTAACAACCTTTCTTTGACCTGCCCCAGAGA 320

Db 242 AATCCAGTGTGGCAAAAAGACTATTATTAACAACCTTTCTTTGACCTGCCCCAGAGA 301

QY 321 GGAGAAATCTGTTGATGAGATTTCTTAAAGATGAAGTGGACAGCCCAAGTCCAGCT 380

Db 302 GGAGAAATCTGTTGATGAGATTTCTTAAAGATGAAGTGGACAGCCCAAGTCCAGCT 361

QY 381 TTCCAGAAAGACAGGAGAAACGGGACAGCCAGCCATTATTCAGACTTACGGGACAC 440

Db 362 TTCCAGAAAGACAGGAGAAACGGGACAGCCAGCCATTATTCAGACTTACGGGACAC 421

QY 441 CCTGGAAGACCAATGCTACCGTGAAGTCCCTACAGAACGCTTTAAACAAGCAGAGAT 500

Db 422 GCTGGAAGACCAATGCTACCGTGAAGTCCCTACAGAACGCTTTAAACAAGCAGAGAT 481

QY	501	GGTGTCTCCACCCTGGAAAAACAGATGAAGTCTCTGAGCAGCGCAGGATGAGACCA	560
Db	482	GCTGTGCTCCACACTGAAAAAGCAGATGAAGTACTTAGACGACGAGGAGATGAGACCA	541.1
QY	561	ACAAAGCTTGGGAGGAGAGGCCACCGACTCAAGTGCAGATGAAAAACCATGGAGCAAAATTGA	620
Db	542	ACAAACACAAAGAGGAGGGCCC - GCGGCTTCAGAGCAGATGAAGCATGGAGCAGATTGA	600
QY	621	GCTCTACTCCAGAGCCAGCGTCTGAGGTGGAGAGATGATTCGAGACATGGGTGTGGC	680
Db	601	GCTTCTACTCCAGAGACCACGCCCCCTTGAGGTGAGAGAAATGATCCGAAACATGGGTGTGGC	660
QY	681	ACAGTCACGCGTGGAGCAGCTGTGTGTACTGTGCGTCTCCCTCAAGAAAAAGTATGAA	740
Db	661	ACAGTCACGCGGTGGAGACCGGTGG - TGTGTACTGTGTCTCTCTCCAGAAAGT - CGAAA	718
QY	741	TCTGAAGGAAGCTGGGAAGC	761
Db	719	TCTAAAAAGACCGGGAGGCC	739

RESULT	16
LOCUS	BG420765
DEFINITION	BG420765 908 bp mRNA linear EST 14-MAR-2001 602445835861 NIH.MGC_14 Homo sapiens CDNA clone IMAGE:4558645 ,
ACCESSION	BG420765
VERSION	BG420765
KEYWORDS	BG420765.1 GI:13327271
SOURCE	EST.
	human.

ORGANISM	Homo sapiens
Eukaryota	Metazoa: Chordata: Craniala: Vertebrata: Euteleostomi:
Mammalia: Eutheria:	Primates: Catarrhini: Hominoidea: Homo.
1 (bases 1 to 908)	
NIH-MGC	http://mgc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)	
Unpublished (1999)	
Contact:	Robert Strausberg, Ph.D.
COMMENT	

Email: cgabps-ft@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LUCM316 row: d column: 15
High quality sequence stop: 719.

```

FEATURES      Location/Qualifiers
source        1. .908

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4586846"
/clone_1ib="NH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOT8; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAC(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using 2AP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
```

Query Match	25.5%	Score 504.2;	DB 10;	Length 908;
Best Local Similarity	82.0%	Pred. No. 4.4e-94;		
Matches 663; Conservative	0;	Indels 93;	Indels 53;	Gaps 5;

91 GCGTCGTTGAGTCGAGCCATCATGCGTATCCCTCTCTCTGTGCACTATCTGCTCCGACTTC 150

[illegible]

RESULT	17
BES46959	
LOCUS	
DEFINITION	BES46959 825 bp mRNA linear EST 09-AUG-2000 601071929P1 NIH_MGC_12 Homo sapiens CDNA clone IMAGE:3458023 5'
ACCESSION	mRNA sequence.
VERSION	BES46959
KEYWORDS	BES46959.1 GI:9775604
SOURCE	EST.
ORGANISM	Homo sapiens human.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukariota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 825)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Inocyte Genomics, Inc.
 DNA Sequencing by: Inocyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM8448 row: j column: 08
 High quality sequence stop: 655.
 Location/Qualifiers

FEATURES

Source

1. 825
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3458023"
 /clone_1lb="NH_MGC_12"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /note="Organ: cervix; Vector: PCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.4 kb. Library prepared by Life
 Technologies."
 BASE COUNT 230 a 209 c 242 g 144 t

ORIGIN

Query Match

Best Local Similarity 85.18; Pred. No. 7.2e-94; Length 825;
 Matches 633; Conservative 0; Mismatches 103; Indels 8; Gaps 6;

QY 135 TATTCGTCCTCCAGCTTTCGATCACTCCCGGAGCTGGCTCCATCCACTGTGGCCACAC 194
 Db 1 TATTCGTCCTCCAGCTTTCGATCACTCCCGGAGCTGGCTCCATCCACTGTGGCCACAC 60

QY 195 TTTTCATCTGCAATGCTTAATCCAGTGGTTTGAACAGCACCACCAAGTGGACCTGCCACA 254
 Db 61 CTTCACCTTGAGTGCCTAATTCAGTGTGAGACAGCACCAAGTGGACCTGCCACA 120

QY 255 GTGTAGATCCAGGTGGCAAAAGACATATTAACAACCTTTCTTGACCTCGCCCA 314
 Db 121 GTGCCGATCCAGGTGGCAAAAGACATATTAACAACCTTTCTTGACCTCGCCCA 180

QY 315 GGAAGAGAGAAATGTTTGATGCAAGATTTTAAAGATGAACAGACAGCTTCAAGC 374
 Db 181 GGAGAGAGAAATGTTTGATGCAAGATTTTAAAGATGAACAGACAGCTTCAAGC 240

QY 375 TCAGCTTCCAGAAAGACAGGAAAGGAGACAGCCGCTTATCGACACTCTACG 434
 Db 241 CCAGCTTCCAGAAAGACAGGAAAGGAGACAGCCGCTTATCGACACTCTACG 300

QY 435 GGACACCTCGAAGAGCAATGCTACCGTGAGTCCCTACAGAAAGCCTTAACAAAGC 494
 Db 301 GGATACGCTGGAAGAGCAATGCTACCGTGAGTCCCTACAGAAAGCCTTAACAAAGC 360

QY 495 AGAGATGCTGTGTTCCACCTGTAAGAAACAGATGAAGTTCTCGAGACAGCGAGATGA 554
 Db 361 CGAGATGCTGTGTTCCACCTGTAAGAAACAGATGAAGTTCTCGAGACAGCGAGATGA 420

QY 555 GACCAACAACTCGGAGAGAGCCCGACGACTCAAGTCAAGATGAAGAAACCATGAGACA 614
 Db 421 GACCAACAACTCGGAGAGAGCCCGACGACTCAAGTCAAGATGAAGAAACCATGAGACA 480

QY 615 AATTGAGCTCTACTCCAGACGAGCGTTTGAGGTGAGAGATGAATTCAGACATGGG 674
 Db 481 GA-TGAGCTTCTACTCCAGACGAGCGCGTTTGAGGTGAGAGATGAATTCAGACATGGG 539

QY 675 TGTGGGACAGTCAAGCGGT-GGAGCAGCTGTGTTTACTGCGTCCCTCAAGAAAGAGT 733
 Db 540 TGTGGGACAGTCAAGCGGTGGGAGAACAGTGTGTTTACTGCGTCCCTCAAGAAAGAGT 599

QY 734 ATGAGAACTCTAAGAGAGCTGGAAAGCC-ACAGGGCAACTGCTGACAGGTTGAAGAG 792
 Db 600 ACAGAAATCTTAAAGAGGACAGGAAAGGCTCCAGGAGAGGTTGCTGACAGGTTGAAGAG 659

QY 793 GATTGGTGTCTCTAGAGCAAGTTGAAGACTCTCAACACTGAGTGAATCAGCCCAAG 852
 Db 660 GA-TGCGTTTCTCTCAGAGGCAAGT---GGACAGTCTACTCCGAATGTGTCAGCCCAAG 715

QY 853 TTGAACTGAGGTCAAGCCCAAG 876
 Db 716 -TAGAACTGAGGTCCGCCCAAGG 738

RESULT 18

AL559248

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

GENOSCOPE

BP 191 91006

EVRV cedex - France

Email: segref@genoscope.cns.fr

Web: www.genoscope.cns.fr

Location/Qualifiers

1. 645

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0D012YH21"

/clone_1lb="LTI_NFL008_TG2"

/sex="male"

/tissue_type="T cells from T cell leukemia"

/note="Vector: pcMVSPORT 6; Site_1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pcMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact: Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax: (1) 301 610 8371

Email: fliang@lifetech.com URL:

<http://fulllength.invitrogen.com>

BASE COUNT 167 a 174 c 177 g 127 t

ORIGIN

Query Match

Best Local Similarity 88.48; Pred. No. 1.2e-92; Length 645;
 Matches 540; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 81 TGGTCCCTGGGCTCTGAGTCCAGGACATGCTTCTCTCTCTCTGTCGATATGTG 140
 Db 35 TGGCTGCTGGGCTCTGAGTCCAGGACATGCTTCTCTCTCTCTGTCGATATGTG 94

QY 141 CTCCGACTTCTTCGATCACTCCCGGAGCTGGTCCATCCACTGTGGCCACACTTTTCA 200
 Db 95 CTCCGACTTCTTCGATCACTCCCGGAGCTGGTCCATCCACTGTGGCCACACTTTTCA 154

QY 201 TCTGCAATGCTTAATCAAGTGTGTTGAGACAGACCAAGTGGAGCTGCCACAGTAG 260
 Db 155 CTGAGAGCTTAATTCAGTGTGTTGAGACAGACCAAGTGGAGCTGCCACAGTAG 214

QY 261 AATCAGGTTGGCAAAAGACTATTAACAACAACTTTTCTTGACCTCGCCAGAGAGA 320
 Db 215 AATCAGGTTGGCAAAAGACTATTAACAACAACTTTTCTTGACCTCGCCAGAGAGA 274

QY 321 GGAGAATGTCTTGATGAGAAATCTTAAAGATGAAGTGAACAGCTCAAGGCTCAGCT 380

|||||
Db 275 GGAGATGTCTTGATGACGATCTCTTAAGATGACATGACATGTCAGAGCCAGCT 334
Oy 361 TTCCCAAGAAAGCAGGAGAAACGCGACAGCCAGCCATTCATGACACTCTACGGGAC 440
Db 335 TTCCCAAGAAAGCAGGAGAAACGCGACAGCCAGCCATTCATGACACTCTACGGGATAC 394
Oy 441 CCTGGAAGAACGCAATGCTACCGTGGAGTCCCTACAGAAAGCCCTTAAACAGCAGAGAT 500
Db 395 GCTGGAAGAACGCAATGCTACCTGTGTGTATCTGTGACAGCAGGCTTGGGCAAGCCGAGAT 454
Oy 501 GCTGTGTTCCACCTGAAAAACAGATGAAGTTCTTGAGAGCAGCGCAGATGAGACCAA 560
Db 455 GCTGTGTTCCACCTGAAAAACAGATGAAGTTAGAGCAGCAGCAGATGAGACCAA 514
Oy 561 ACAACCTGGGAGAGGCGCCACCGACTCAAGTCAGATGAAGAACCATGAGACAAATGCA 620
Db 515 ACAACCAAGAGAGAGGCGCCCGGCTCAGAGCAAGATGAAGAACCATGAGACAAATGCA 574
Oy 621 GCTCCTACTCCAGAGCCAGCGCTTCTGAGGTGAGAGATGATTCGAGACATGGTGTGG 680
Db 575 GCTTCTACTCCAGAGCCAGCGCCCTGAGGTGAGAGATGATCCGAGACATGGTGTGG 634
Oy 681 ACAGTCACCGG 691
Db 635 ACAGTCACCGG 645

RESULT 19
BE869186 888 bp mRNA linear EST 20-OCT-2000
LOCUS 601445180P1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3849451 5',
DEFINITION mRNA sequence.
ACCESSION BE869186
VERSION BE869186.1 GI:10317962
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 888)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov

Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LHAM9566 row: O column: 20
High quality sequence stop: 692.

FEATURES
source
1..888
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3849451"
/clone_id="NIH_MGC_65"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site: 1; NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies."

BASE COUNT 254 a 199 c 268 g 167 t
ORIGIN

Query Match 24.8%; Score 490.6; DB 10; Length 888;
Best Local Similarity 79.8%; Pred. No. 2.9e-91;

Matches 665; Conservative 0; Mismatches 159; Indels 9; Gaps 7;
Oy 319 GAGGAGATGCTCTTGATGACGATTTCTTAA--GAATGAACTGGACACGCTCAAGCTCA 377
Db 8 GAGGAGATGCTCTTGATGACGATTTCTTAAAGGATGAATGACATGTCAGAGCCCA 67
Oy 378 GCTTTCAGAAAGCAGGAGAAACGCGACAGCCAGCCATTCATGACACTCTACGGGA 437
Db 68 GCTTTCAGAAAGCAGGAGAAACGAGACAGCCAGGTCATCATGACACTCTGCGGGA 127
Oy 438 CACCTGGAAGAGCGC-AATGCTACCGTGAAGTCCCTACAGAACGCTTAAACAGGAC 496
Db 128 TACGCTGGAAGAGCGCGAATGCTACTGTGTATCTCTCAGACAGGCTTGGGACGCG 187
Oy 497 AGATGCTGTGTTCCACCCCTGAAAAACAGATGAAGTTCTTGAGAGCGGAGAGATGGA 556
Db 188 AGATGCTGTGTTCCACACTGAAAAAGCAGATGAAGTACTTGAAGCAGCAGAGATGGA 247
Oy 557 CCNAAAGCTCGGAGAGAGGCCACCGACTCAAGTCAGATGAAGTGAAGCAATGAGCAAA 616
Db 248 CCNAAAGCAGCAGAGAGAGGCCCGCGGCTCAGAGAGCAAGATGAAGCAATGAGCA 307
Oy 617 TTGAGCTCTTACTCCAGAGCCAGCGCTTGAAGTGGAGAGATGATTCGAGACATGGGT 676
Db 308 TTGAGCTCTTACTCCAGAGCCAGCGCCTGAGGTGGAGAGATGATTCGAGACATGGGT 367
Oy 677 TGGGACAGTCAGCGGTGGAGAGCGGTGTGTACTGTGCTCCCTCAAGAAAGATGTG 736
Db 368 TGGGACAGTCAGCGGTGGAGAGCGGTGTGTACTGTGTCTCTCAAGAAAGATGCG 427
Oy 737 AGAATCTGAAGGAAGCTCGGAGAGGCCACAGGGAAGTGGTACAGAGTTGGAAGAGATT 796
Db 428 AGAATCTGAAGGAAGCGACGGAAGCGCTCAGGGAGGTGGTGAACAGCTGAGGAAGA -T 486
Oy 797 TGTGTCTCTTAGAGAGCAAGTTGAAGACTCTCAACACTGAGCTGATGAGCCAAATTAG 856
Db 487 TGTGTCTCTTAGAGAGCAAGTTGAAGACTCTCAACACTGAGCTGATGAGCCAAATTAG 546
Oy 857 AACTGAGGTGACGCCAGAGAGCACTTACAAAGTGTGACACAGAGATGACGAGCTTAAGA 916
Db 547 AACTGAGTCAAGCCAGAGAGCACTTACAGAGTGTGACAAAGAAATCATGAGCTGAAGA 606
Oy 917 AGAATCTGATGATCTCTCAGAGGAACCTTGAGCC--TGCCCTCCGCGACCAATGAGACGG 973
Db 607 AGAATCTGAAGAGATGCTCAGAGAAACCTTGACCTGTCCACAGGTGCGAGGTAGACTG 666
Oy 974 TCAGCCGCTGTGTTTGAAGAGCCAGCCCTGT--GAGATGATGAACCGAGGCTTCAC 1032
Db 667 TCAGCCGCTGTGTTTGAAGAGCCAGCCCTGTGAGGAGTGAATCTGGAAGGCTCGGCC 726
Oy 1033 CAGCCACCTTCGGTGAAGAGAT--TGATCTCAATACCACTTGATGTAATACCCCTCC 1091
Db 727 GGGCATCTATCTCCGAGATGATGTGATCAATGCTTAATGTTGGGTATATC-CC 785
Oy 1092 AACCCAGACTCTGTGCTCCAGACTTGGCTCCCAAGAGAGCTGAGCTCGGAGA 1144
Db 786 CAGCCGCGCTCAGCTTCCAGCATGGGTACTAGAAAACCTTGCTGAGAGAGA 838

RESULT 20
B1758046 668 bp mRNA linear EST 25-SEP-2001
LOCUS 603030577P1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200986 5',
DEFINITION mRNA sequence.
ACCESSION B1758046
VERSION B1758046.1 GI:15749624
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 668)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

Db 303 GGAGATCTCTTGATGCAATTCCTTAAGATGAAGTGGACAATGTGACAGCCCACT 362
QY 381 TTCCAGAAAGCAGAGGAAACGGACAGCCAGCCATTATCGACACTTACGGACAC 440
Db 383 TTCCAGAAAGCAGAGGAAACGGACAGCCAGCCATTATCGACACTTACGGACATC 422
QY 441 CCTGGAGACGCAATGCTTACCGTGGAGTCCCTTACAGAGCCCTTAAACAAGCAGAGAT 500
Db 423 GCTGAAACAGCAGCAATGCTTACCGTGGAGTCCCTTACAGAGCCCTTAAACAAGCAGAGAT 482
QY 501 GCTGTTTCCACCCGAAAGAAAGATGAGTCTCTGGAGAGCCGGCAGATGAGACCAA 560
Db 483 GCTGCTCAGACAGTAAAGAAAGCAGATGAGTCTTAAAGAGCAGAGATGAGACCAA 542
QY 561 ACAAGCTGGAGAGAGGCCACCGACTCAAGTGCAGAGTGAAGAAACATGAGACCAATTGA 620
Db 543 ACAAGCACAAGAGAGAGGCCCGCCGCTCAGAGAGCAAGTGAAGAACATGAGACCAATTGA 602
QY 621 GCTCCTACTCCAGAGCCAGCCTTCTGAGAGTGGAGAGATGATTCGAG-ACATGGGTGG 679
Db 603 GCTTCTACTCCAGAGCCAGCG-CCGTGAGTGGAGAGATGATTCGAGAACATGGGTGG 661
QY 680 GACAGTCGCGGTGGAGAGCTGGCTGTGT-ACCTGCTGTCCCTCAAGAAAGATATGAG 738
Db 662 GACAGTCAGCGGGGAAACCGTGGCTGTGTAAAGGGGGTCTCCAGAGAAAGCCTCCAA 721
QY 739 AATCTGAAGAGAGCTCGGAGGCCA 763
Db 722 AATCCAAACAGAGGCCGAGAGCTCA 746

RESULT 22
BF219707 632 bp mRNA linear EST 08-NOV-2000
LOCUS 601296627F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2821007 5',
DEFINITION mRNA sequence.
ACCESSION BF219707
VERSION BF219707.1 GI:11125801
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 632)
NIH-MGC <http://mgc.nci.nih.gov/>.
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabps-remail.nih.gov
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: IRAL1 row: e column: 07
High quality sequence stop: 630.
Location/Qualifiers
1. 632
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2821007"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site: 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

FEATURES
source

BASE COUNT 163 a 171 c 172 g 126 t
ORIGIN
Query Match 24.1%; Score 476.2; DB 10; Length 632;
Best Local Similarity 87.7%; Pred. No. 2.8e-88;
Matches 520; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 81 TGGTTCCTGGGCGCTTGAGTCCAGCCATCATGCTATCCCTCTGTGCTACTATCTG 140
Db 37 TGGCTGCTGGGCCCCCTTGAGTCCAGCCATCATGCTATCCCTCTGTGCTACTATCTG 96
QY 141 CTCGACTTCTTCGATCACTCCCGTGAAGTGGTCCCATCCATGTCGCGACACTTTTCA 200
Db 97 CTCGAGCTTCTTCGATCACTCCCGGAGAGTGGCGCCCATCCATGTCGCGACACTTTCA 156
QY 201 TCTGCAATGCTTAATTCAGTGGTTTGAGACAGCACAAGTGGACCTGCGCACAGTGTAG 260
Db 157 CTTCAGTGGCTTAATTCAGTGGTTTGAGACAGCACAAGTGGACCTGCGCACAGTGG 216
QY 261 AATCAGGTTGGCAAAAAGACTATTATTAACAACCTTTCTTGACCTGCGCCAGAGAGA 320
Db 217 AATCCAGGTTGGCAAAAAGACTATTATTAACAACCTTTCTTGATCTTGGCCAGAGAGA 276
QY 321 GGAGAATGCTTTGGATGCGAGATTTCTTAAAGATGAAGTGGACAGCCTCAAGCTCAGCT 380
Db 277 GGAGAATGCTTTGGATGCGAGATTTCTTAAAGATGAAGTGGACAGCCTCAGCTCAGCT 336
QY 381 TTCCAGAAAGCAGAGGAAACGGACAGCCAGCCATTATGACACTTACGGGACAC 440
Db 337 TTCCAGAAAGCAGAGGAAACGGACAGCCAGCCATTATGACACTTACGGGAGTAC 396
QY 441 CCTGGAAGAACGCAATGCTATACCGTGGAGTCCCTCAAGACGCTTAAACAAGCAGAGAT 500
Db 397 GCTGGAAGAACGCAATGCTATACCGTGGAGTCCCTCAAGACGCTTGGGCAAGCCGAGAT 456
QY 501 GCTGTGTTCCACCCCTGAAGAAACAGATGAAGTTCTTGAGAGCAGCGCAGAGATGACCAA 560
Db 457 GCTGTGCTCAGCACTGTAAGAAAGAGATGAAGTACTTGAAGAGCAGCAGATGAGACCAA 516
QY 561 ACAAGCTCGGAGAGAGGCCACCGACTCAAGTGCAGAGTGAAGAAACATGAGCAATTTGA 620
Db 517 ACAAGCACAAGAGAGAGGCCCGCCGCTCAGAGAGCAAGATGAAGCACTGAGAGATTTGA 576
QY 621 GCTCCTCTCCAGAGCCAGCGCTTCTGAGAGTGGAGAGATGTTGGAGACATGG 673
Db 577 GCTTCTACTCCAGAGCCAGCGCCCTGAGGTGGAGAGATGATCCGAGACTGGG 629

RESULT 23
BG117979 912 bp mRNA linear EST 30-JAN-2001
LOCUS 602351223F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4446149 5',
DEFINITION mRNA sequence.
ACCESSION BG117979
VERSION BG117979.1 GI:12611485
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 912)
NIH-MGC <http://mgc.nci.nih.gov/>.
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabps-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.lnl.gov
 Plate: L1AM10224 row: n column: 06
 High quality sequence stop: 678.
 Location/Qualifiers

FEATURES

source

1.912
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4446149"
 /clone_lib="NIH.MGC.90"
 /tissue_type="adenocarcinoma, cell_line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Liver; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH.MGC Library."

BASE COUNT 236 a 243 c 273 g 160 t
 ORIGIN

Query Match 23.3%; Score 459.8; DB 10; Length 912;
 Best Local Similarity 83.8%; Pred. No. 6.0e-85;

Matches 591; Conservative 0; Mismatches 107; Indels 7; Gaps 6;

OY 81 TGGTTCCTGGGCTGCTGAGTCGAGCCATCATGCTATCCCTCTCTGTCGACTATCTG 140
 ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 65 TGGCTGCTGGGCCCCCTTGAATCCAGCCATCATGCTATCCGCTCTGTCGACTATCTG 124
 OY 141 CTCGACCTCTTGGATCACTCCCGTGAAGTGGCTGCCATCCATCGTGGCCACACTTTTCA 200
 ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 125 CTCGACCTCTTGGATCACTCCCGGAGAGTGGCCGATCCATCGGCGCCACACTTTTCA 184
 OY 201 TCTGCATTCCTATTCAGTGTGAGACAGCAACAGTGGAGCTGCCACAGTGTAG 260
 ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 185 CTTCAGTGCCTAATTAAGTGTGAGACAGCAACAGTGGAGCTGCCACAGTGTAG 244
 OY 261 AATCAGGTTGGCAAAAGAACTATTATTAACAACTTTCTTGAAGCTGGCCAGAGAGA 320
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 245 AATCAGGTTGGCAAAAGAACTATTATTAACAACTTTCTTGAAGCTGGCCAGAGAGA 304
 OY 321 GGAGATGCTTGGATGCAAGATTTCTTAAGATGAAGTGCAGAGCTCAAGCTCA-GC 379
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 305 GGAGATGCTTGGATGCAAGATTTCTTAAGATGAAGTGCAGAGCTCAAGAGCTCAAG 364
 OY 380 TTTCGCAAG 439
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 365 TTTCGCAAG 424
 OY 440 CCTGGAAG 498
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 425 CGCTGGAAG 484
 OY 499 ATGCTGTCTCCACCTGAAAAAAGATGAGTCTCTGAGACAGGCGAGAGATGAGACC 558
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 485 ATGCTGTCTCCACCTGAAAAAAGATGAGTCTCTGAGACAGGCGAGAGATGAGACC 544
 OY 559 AAAAAGGTGGGAG 618
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 545 AAAAAGGTGGGAG 604
 OY 619 GAGCTCTCTACT-CCAGAGCCAGAGTCTTGAAGTGGAGAGAG-ATGATTTGAGACAT-GGGT 675
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 605 GAGCTTCTACTCCAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 664
 OY 676 GTGGAGACGTAGCGGT--GGAGAGCTGCTGTCTACTGCTGCTCTCAAGAAAGAGT 733
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 665 GTGGAGACGTAGCGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 724
 OY 734 ATGGAATCTGAAGAGAGCTGGAAGGCGACAGGGGAGAGAGAGAGAGAGAGAGAGAG 778
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 725 ACGACATCTACAG 769

RESULT 24
 BB821649
 LOCUS

BB821649 RIKEN full-length enriched, mammary gland RCB-0526
 Jy9-MC(A) CDNA Mus musculus cDNA clone G830012P09 3', mRNA

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

Unpublished (2001)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh

, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)

wag1, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and
 Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multiplexed sequencer. Genome Res. 10
 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
 Y., and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.
 e mouse tissues.

FEATURES

source

1.528
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="G830012P09"
 /clone_lib="RIKEN full-length enriched, mammary gland
 RCB-0526 Jy9-MC(A) CDNA"
 /tissue_type="mammary gland"
 /cell_line="RCB-0526 Jy9-MC(A)"

BASE COUNT

ORIGIN

Query Match

22.1%; Score 437.2; DB 9; Length 528;
 Best Local Similarity 93.2%; Pred. No. 3.3e-80;

Matches 490; Conservative 0; Mismatches 33; Indels 3; Gaps 3;

OY 1434 TAACTCAAGGCCAAGAGTAACAGAAAGTGAATTAAGACTGTGATTCCTGCCCTCCA 1493
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 2 TAAGTCCAAAGGCCAAGAAATTAACAGAAATTGAGAAATTAAGACTGTAGTTCCTGCCCA 61
Oy 1494 -GCCCAAGCTGGATACCTTCTTATGTAGTGAAGCGGTGACCGAGTGAATGTTTGCATTA 1552
Db 62 GGCCCAAGCTGGATACCTTCTTATGTAGTGAAGCGGTGACCGAGTGAATGTTTGCATTA 121
Oy 1553 GTGGGCCAAGACCTGGCTTAACCGGAAGTGTGTTTGAAGATGGCTCTTGTGACAGTTC 1612
Db 122 GTGGGCCAAGACCTGGCTTAACCGGAATGTTTGAAGATGGCTCTTGTGACAGTTC 181
Oy 1613 CAAGAGATATGCCCAAGAAACACACTTCTGTTGTTCACTGGG -CCCTGACACACACTGGG 1671
Db 182 CAAGGAATATGCCCAAGAAACACACTTCTGTTGTTCACTGGG -CCCTGACACACACTGGG 241
Oy 1672 AAGCCACATGACCACTGTTACTGTTCCGATCAG -CAGGCGCTACTCCAGTTGACAGGTTT 1730
Db 242 AAGCCACATGACCACTGTTACTGTTCCGATCAG -CAGGCGCTACTCCAGTTGACAGGTTT 301
Oy 1731 TGCCTATAGCTAACAACAGAGTGTGCTGACTCTTTTGTGTTTATAGAACAGGGTCA 1790
Db 302 TGCCTATAGCTAACAACAGAGTGTGCTGACTCTTTTGTGTTTATAGAACAGGGTCA 361
Oy 1791 TTGACTTAAGTGATGGAGTGGAGTGGAGATCCTATGACAGGCTGAGACCTGCGCTT 1850
Db 362 TTGACTTAAGTGATGGAGTGGAGTGGAGATCCTATGACAGGCTGAGACCTGCGCTT 421
Oy 1851 GAACCTCTGCTGCTGCTCCAGCTTATGCTTGAATTAATGAGTGAAGTGAAGGAA 1910
Db 422 AAACCTCTGCTGCTGCTCCAGCTTATGCTTGAATTAATGAGTGAAGTGAAGGAA 481
Oy 1911 AGCTTGGGGAAGTGTGCTGTAATAAATGAAGGATCTTTTCTTC 1956
Db 482 AGCTTGGGGAAGTGTGCTTAAATAAATGAAGGATCTTTTCTTC 527

RESULT 25
AM490717/c 441 bp mRNA linear EST 24-FEB-2000
LOCUS UI-M-BH3-asx-c-07-0-UI.s1 NIH_BMAP_M.S4 Mus musculus cDNA clone
DEFINITION UI-M-BH3-asx-c-07-0-UI 3', mRNA sequence.
ACCESSION AM490717
VERSION AM490717.1 GI:7060988
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 441)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Chln, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mstt@mail.nih.gov

The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the Molt site
and the oligo-dT track served to identify it as a clone from the
normalized prefrontal cortex library cDNA library Preparation: M.B.
Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
clones from RESEARCH GENERICS. It should be noted that Bento Soares
is generating a small number of additional specialized
non-redundant arrays of BMAP cDNAs whose availability will be
considered under appropriate and limited collaborative arrangements
Seq primer: M13 Forward
POLYA=yes.

FEATURES
source

Location/Qualifiers
1..441
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH3-asx-c-07-0-UI"
/clone_1id="NIH_BMAP_M.S4"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; The
NIH_BMAP_M.S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH_BMAP_M.S4,
NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, NIH_BMAP_M.S3.1,
NIH_BMAP_M.S2, NIH_BMAP_M.S1. The subtracted library
(NIH_BMAP_M.S4) was constructed as follows: PCR amplified
cDNA inserts from NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and
NIH_BMAP_M.S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and NIH_BMAP_M.S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
NIH_BMAP_M.S4 library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)
TAG_LIB=NIH_BMAP_M.S4
TAG_TISSUE=prefrontal.cortex
TAG_SEQ=GCCTCA"

BASE COUNT 117 a 115 c 94 g 115 t
ORIGIN

Query Match 21.9%; Score 433; DB 9; Length 441;
Best Local Similarity 98.9%; Pred. No. 2.5e-79;
Matches 436; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1534 AGAGTGATGTTTGGCAATTAATGAGGCGCAAGACCTGGCTAACCAGGAGTGTGGAAGAT 1593
Db 441 AGAGTGATGTTTGGCAATTAATGAGGCGCAAGACCTGGCTAACCAGGAGTGTGGAAGAT 382
Oy 1594 GGCTCGCTTGGAGCAGGCGCAAGAGAGTGGCGCCAGAAACACACTTCTGTTCACTGC 1653
Db 381 GGCTCGCTTGGAGCAGGCGCGCAAGAGAGTGGCGCCAGAAACACACTTCTGTTCACTGC 322
Oy 1654 GCCCTGCACACACACTGGGAAGCCACATGACAGTTTACTGTTCCGATCAGACAGGCGCTAC 1713
Db 321 GCCCTGCACACACACTGGGAAGCCACATGACAGTTTACTGTTCCGATCAGACAGGCGCTAC 262
Oy 1714 TTCCAGTTGCAAGGTTTGTCTTATAGCTACACACAGGTGTGCTGACCTCTTTGTTT 1773
Db 261 TTCCAGTTGCAAGGTTTGTCTTATAGCTACACACAGGTGTGCTGACCTCTTTGTTT 202
Oy 1774 TATGAACAGGCGTACACTACCTAGTGGATGGAGTGGAGATGCTGAGATGCTATGCAAGC 1833
Db 201 TATGAACAGGCGTACACTACCTAGTGGATGGAGTGGAGATGCTGAGATGCTATGCAAGC 142
Oy 1834 TGGAGACCTGCGCTTGAACCTCTGCTGCTCCACACTTATGCTGAATTAATGAGGAT 1893
Db 141 TGGAGACCTGCGCTTGAACCTCTGCTGCTCCACACTTATGCTGAATTAATGAGGAT 82
Oy 1894 GAGGTGTGATAGGGAAGGTTGGGGAAGTTCGTGTGAATAAATGAAGGATCTTTTC 1953
Db 81 GAGGTGTGATAGGGAAGGCTGGGGAAGTTCGTGTGAATAAATGAAGGATCTTTTC 22

Db 459 GGAGGACGAACAAATTCATCCACCTAGGAGACACCACTTATCCAGCAGTCCCTGT 400
 QY 1435 AAGTCCAGGCGCAGAGTAAACAGAAAGTGAAGTAAAGACTGTGAGTTCGCTCCAG 1494
 Db 399 AAGTCCAGGCGCAGAGTAAACAGAAAGTGAAGTAAAGACTGTGAGTTCGCTCCAG 340
 QY 1495 CCCAAGCTGATACCTTCTATGTCACTGAAAGCTGACAGAGTATGTTGCAATTAGT 1554
 Db 339 CCCAAGCTGATACCTTCTATGTCACTGAAAGCTGACAGAGTATGTTGCAATTAGT 280
 QY 1555 GGGCCAGAGCTGCTTACCGGAAGTCTTTTGGAGAATGCTCTCTGGACCAATCCA 1614
 Db 279 GGACCAAGCTGCTTACCGGAAGTCTTTTGGAGAATGCTCTCTGGACCAATCCA 222
 QY 1615 AGAGAGATGCCCAAAACACACTTCTGCTGACGCGCCGCGCCACACAGGGGAG 1674
 Db 223 -----TTCTGTGTTCACTGCGCCGCGCCACACAGGGGAG 187
 QY 1675 CCACATGACAGTCTTACTGTTCCGATCAGAGGCTACTTCCAGTTCAGAGGTTTGTCT 1734
 Db 186 CCAGGTGACAGTCTTACTGTTCCGATCAGAGGCTACTTCCAGTTCAGAGGTTTGTCT 127
 QY 1735 TATAGTACAAACAGGTGTGCTGACTCCTTTGTTTATAGAACAGGCTCAATTGA 1794
 Db 126 TATAGTACAAACAGGTGTGCTGACTCCTTTGTTTATAGAACAGGCTCAATTGA 67
 QY 1795 CTCAAGTGAATGGAGGCTGGAGGATCCTATGACAGGCTGAGAGGCTTCCGCTTGAAC 1854
 Db 66 CTCTCAGTGAATGGAGGCTGGAGGATCCTATGACAGGCTGAGAGGCTTCCGCTTGAAC 8

RESULT 29
 BE031265 547 bp mRNA linear EST 09-JUL-2000
 LOCUS
 DEFINITION 129770 MARC 1P16 Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION BE031265
 VERSION BE031265.1 GI:8326274
 KEYWORDS EST.
 SOURCE pig.
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 547)
 Fahrenkrug,S.C., Fekking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
 Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.M.,
 and Keele,J.W.
 Design and use of two pooled tissue normalized cDNA libraries for
 EST discovery in swine
 Unpublished (2000)
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt. trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -mismatch 12 options.
 PCR Primers
 FORWARD: AGGAACACGCTATGACCAT
 BACKWARD: GTTTCCTCAGTCACGACG
 Plate: 65 row: H column: 4
 Seq primer: ATTAGGTGACACATATG.
 Location/Qualifiers
 1..347
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_lib="MARC 1P16"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 library made from pooled tissue from day 11, 13, 15, 20,
 and 30 embryos."

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 SOURCE

1..347
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_lib="MARC 1P16"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
 library made from pooled tissue from day 11, 13, 15, 20,
 and 30 embryos."

BASE COUNT 166 a 130 c 161 g 90 t
 ORIGIN
 Query Match 21.5%; Score 424.4; DB 9; Length 547;
 Best Local Similarity 86.1%; Pred. No. 1,4e-77;
 Matches 470; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 219 GTGGTTGAGACACACACAGTGGACCTGCCCACTGTATGATCCAGGTGGCAAAA 278
 Db 2 GTGGTTGAGACACACAGTGGACCTGCCCACTGTATGATCCAGGTGGCAAAA 61
 QY 279 GACTATTATTAACAACTTTCTTGTGACTCGCCAGAGAGAGAGATGCTTGTGATGC 338
 Db 62 AACATTATCAATTAAGCTCTTCTTGTGACTCGCCAGAGAGAGAGATGCTTGTGATGC 121
 QY 339 AGAATCTTAAAGAAATGAACTGACAGCGTCAAGCTTCCAGAAAGACAGGA 398
 Db 122 AGAATCTTAAAGAAATGAACTGATTAATACAGAGCCCTGCTTCCAGAAAGAGGA 181
 QY 399 GAAAGGAGACAGCAGCGCCATATGACACTCTACGAGGACACCTGGAAAGACCAATGC 458
 Db 182 AAAAGAGACAGCAGCGCCATATGACACTCTACGAGGACACCTGGAAAGACCAATGC 241
 QY 459 TACCGTGAAGTCCCTACAGAAACGCTTAAACAGAGAGAGTGTGTTCCACCTGAA 518
 Db 242 CACTGTGGAATCCCTGAGAAAGGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301
 QY 519 AAAACAGATGAATGTTCTGAGAGCGGAGAGATGAGACCAACAGCTGGAGAGGC 578
 Db 302 GAAACAGATGAATGTTCTGAGAGCGGAGAGATGAGACCAACAGCTGGAGAGGC 361
 QY 579 CCACGACCTCAAGTCAAGTGAAGAAACATGAGCAAAATGAGTCTTACTCCAGAGCA 638
 Db 362 CGCCGAGCTCGGAGCAAGATGAGAGACTATGAGAGAGAGAGTCTTACTCCAGAGCA 421
 QY 639 GCGTCTGAGGTGAGAGAGATGATTCGACAGATGGGTGTGGAGACAGTCCAGAGCA 698
 Db 422 GCGGCTGAGGTGAGAGAGATGATTCGACAGATGGGTGTGGAGACAGTCCAGAGCA 481
 QY 699 GCTGGCTGTACTGCGTGTCTCCCAAGAAAGATGAGATCTGAGAGAGCTGGAA 758
 Db 482 GCTGGCTGTACTGCGTGTCTCCCAAGAAAGATGAGATCTGAGAGAGCTGGAA 541
 QY 759 GCGCCAC 764
 Db 542 GGCTTC 547

RESULT 30
 BE031869 450 bp mRNA linear EST 19-NOV-2001
 LOCUS
 DEFINITION BB031869 RIKEN full-length enriched, mammary gland RCB-0527
 JY9-MC(B) cDNA Mus musculus clone G930019H09 3', mRNA
 ACCESSION BE031869
 VERSION BE031869
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclerogasthi; Muridae; Murinae; Mus.
 1 (bases 1 to 450)

REFERENCE
 AUTHORS
 TITLE
 Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
 Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K., Ishii,
 Y., Ito,M., Kawai,J., Kojima,Y., Kono,H., Kouda,M., Matsuyama,T.,
 Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
 Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
 Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,
 A., Takahashi,F., Takaku-Akenita,S., Tanaka,F., Tomaru,A., Toyota,T.,
 Watanabe,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
 2001)

JOURNAL
COMMENT

Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura
, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

FEATURES

Location/Qualifiers
1. 450

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_id="RIKEN full-length enriched, mammary gland
RCB-0527 Jyg-MC(B) cDNA"
/tissue_type="mammary gland"
/cell_line="RCB-0527 Jyg-MC(B)"
BASE COUNT 105 a 102 c 114 g 129 t
ORIGIN

Query Match 21.4%; Score 422; DB 9; Length 450;
Best Local Similarity 97.6%; Pred. No. 4.6e-77;
Matches 439; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

OY 1507 ACCTTCTATGTCACTGAACGGTGACCAAGTGTTCATTTAGTGGGCCAAGACCT 1566
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 1 ACCCTTTATGTCACTGAACGGTGACCAAGTGTTCATTTAGTGGGCCAAGACCT 60
OY 1567 GCGTAAACGGAGATGTTTTTGAAGATGGCTCCTTGACACAGCCCAAGAGAGATGGCC 1626
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 61 GCGTAAACGGAGATGTTTTTGAAGATGGCTCCTTGACACAGCCCAAGAGAGATGGCC 120
OY 1627 AGAAACACACTTCTGTGTTCACATGCGCCCTGCACACACTGGGAAGCCACATGACAG 1686
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 121 AGAAACACACTTCTGTGTTCACATGCGCCCTGCACACACTGGGAAGCCACATGACAG 180
OY 1687 TTTACTGTTCCGATCAGCAGGGCTACTTCCAGTTGACAGGGTTTGCCTTATAGCTACAC 1746
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 181 TTTACTGTTCCGATCAGCAGGGCTACTTCCAGTTGACAGGGTTTGCCTTATAGCTACAC 240
OY 1747 CAGGTGTGGCGGACTCTTTTGTATTATAGAACAGGGTCACATGACTTAAGTGAT 1806
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 241 CAGGTGTGGCGGACTCTTTTGTATTATAGAACAGGGTCACATGACTTAAGTGAT 299
OY 1807 GGAAGTGTGGAGATCTTATGACAGGCTGAGAGACCTCGCTTGAATCCTGCTGCT 1866
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 300 GGAAGTGTGGAGATCTTATGACAGGCTGAGAGACCTCGCTTGAATCCTGCTGCT 359
OY 1867 CCAGCTATGCTTGAATTAAGGGTGAGGTGATAGGAAAGGTTGGGGAAGTTT 1926
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

DB 360 CCAGCTATGCTTGAATTAAGGCGGTGATGATGAGCAAGTGGGAGATT 419
OY 1927 CTGCTAAATAAAGGATCTTTCTTC 1956
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 420 CTGCTAAATAAAGGATCTTTCTTC 449

RESULT 31

BB830067 453 bp mRNA linear EST 19-NOV-2001
LOCUS BB830067 RIKEN full-length enriched, mammary gland RCB-0527
DEFINITION Jyg-MC(B) cDNA Mus musculus cDNA clone G930010B01 3', mRNA
sequence.

ACCESSION BB830067
VERSION BB830067.1 GI:17008310

KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota: Chordata: Cniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
1 (bases 1 to 453)

AUTHORS Akimura, T., Aikawa, T., Carninci, P., Furuno, M., Hangaki, T.,
Hayatsu, N., Hirokawa, T., Hirokawa, T., Imotani, K., Ishii
, Y., Ito, M., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Matsuyama, T.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T.,
Saito, R., Sakai, C., Sakai, R., Sakazume, N., Sasaki, D., Sato, K.,
Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa
, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyoi, T.,
Watabiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.
2001)

TITLE

JOURNAL

COMMENT

Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura
, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

FEATURES

SOURCE

Location/Qualifiers
1. 453
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_id="RIKEN full-length enriched, mammary gland
RCB-0527 Jyg-MC(B) cDNA"
/tissue_type="mammary gland"
/cell_line="RCB-0527 Jyg-MC(B)"
BASE COUNT 104 a 99 c 119 g 131 t
ORIGIN

SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 423)
AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imctani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Kono,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watanishi,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
Riken Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001).
TITLE Unpublished (2001)
JOURNAL Contact: Yoshihide Hayashizaki
COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagii,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanishi,M., Yoneda,Y., Ishikawa,T., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.
FEATURES
location/Qualifiers
1..423
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="G430042L20"
/clone_1lb="RIKEN full-length enriched, Nullipotent stem cell CRT-2070 NE cDNA"
/cell_type="Nullipotent stem cell"
/cell_line="CRL-2070 NE"
BASE COUNT 100 a 93 c 111 g 119 t
ORIGIN
Query Match 20.3%; Score 401.4; DB 9; Length 423;
Best Local Similarity 98.3%; Pred. No. 8.4e-73;
Matches 416; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 1534 AGAGGATGTTTGCATATAGTGGCCAGACCTGGCTAACCGGAAGTGTGGGAAGAT 1593
|||||
DB 1 AGAGCATGTTTGCATATAGTGGCCAGACCTGGCTAACCGGAAGTGTGGGAAGAT 59
QY 1594 GGTCTCTCTGGACAGTCCAGAGAGATGCCAGAAAAACACACTCTCTGTGTTCATGC 1653
|||||
DB 60 GATCTCTCTGGACAGTCCAGAGAGATGCCAGAAAAACACACTCTCTGTGTTCATGC 119
QY 1654 GCCCTGCACACATGGGAAGCCACATGACATTTACTGTTCGATGAGCAGGCGCTAC 1713
|||||
DB 120 GCCCTGCACACATGGGAAGCCACATGACATTTACTGTTCGATGAGCAGGCGCTAC 179

QY 1714 TTCAGTTGCAGGCTTTTGTCTATAGCTACACAGGAGTGTGGCTGGACTCCCTTTGTTT 1773
|||||
DB 180 TTCAGTTGCAGGCTTTTGTCTATAGCTACACAGGAGTGTGGCTGGACTCCCTTTGTTT 239
QY 1774 TATAGACAGGCTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1833
|||||
DB 240 TATAGACAGGCTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 299
QY 1834 TGGAGACCTTCCTGCTTGAACCTTCCTGCTTCACCTTATGCTTAAATATATGGGT 1893
|||||
DB 300 TGGAGACCTTCCTGCTTGAACCTTCCTGCTTCACCTTATGCTTAAATATATGGGT 359
QY 1894 GAGTGTGATAGGGAAGGTTGGGAGTGTTCGTTAAATTAAGGATCTTTTC 1953
|||||
DB 360 GAGTGTGATAGGGAAGGTTGGGAGTGTTCGTTAAATTAAGGATCTTTTC 419
QY 1954 TTC 1956
|||
DB 420 TTC 422
RESULT 34
A1428513 403 bp mRNA linear EST 09-MAR-1999
LOCUS
DEFINITION mo37h05.x1 life Tech mouse embryo 13 5dpc 10666014 Mus musculus
ACCESSION A1428513
VERSION A1428513.1 GI:4274439
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 403)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harrey,N., Schurk,R., Riltter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
CONTACT: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
This clone was previously sequenced on the 5' end only, this new data is from the 3' end
Possible reversed clone: polyT not found
High quality sequence stop: 402.
location/Qualifiers
1..403
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:555801"
/clone_1lb="Life Tech mouse embryo 13 5dpc 10666014"
/tissue_type="embryo"
/dev_stage="13 5dpc embryos"
/lab_host="DH10B"
/note="Organ: whole embryo; Vector: pCMV-SPORT2; Site_1: Salt; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. 13.5dpc embryos. pCMV-SPORT2 vector."
BASE COUNT 110 a 108 c 88 g 97 t
ORIGIN
Query Match 20.2%; Score 399.8; DB 9; Length 403;
Best Local Similarity 99.5%; Pred. No. 1.8e-72;

Matches 401; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1560 AAGACCTGCTTAACCGAAGTGTGTTTGAAGATGAGTCTCTTGGACAGTCCAAAGAGA 1619
|||||
Db 403 AAGACCTGCTTAACCGAAGTGTGTTTGAAGATGAGTCTCTTGGACAGTCCAAAGAGA 344
QY 1620 GATGCCAGAAAACACACTCTCTGTGTCTACCTGCGCCCTGACACACACTGGGAAAGCCACA 1679
|||||
Db 343 GATGCCAGAAAACACACTCTCTGTGTCTACCTGCGCCCTGACACACACTGGGAAAGCCACA 284
QY 1680 TGACACAGTTTACTGTTCGGATCAGAGGCGCTACTCTCCAGTTGCAAGGTTTGGCTTATAG 1739
|||||
Db 283 TGACACAGTTTACTGTTCGGATCAGAGGCGCTACTCTCCAGTTGCAAGGTTTGGCTTATAG 224
QY 1740 CTACACACAGGTGTGCTGAGACTCTTGTGTTTATAGAAGAGGTACATTCAGCTTA 1799
|||||
Db 223 CTACACACAGGTGTGCTGAGACTCTTGTGTTTATAGAAGAGGTACATTCAGCTTA 164
QY 1800 AGTGGATGGAGTGTCTGAGAGATCCTATGACAGCTGAGAGACCTGCGCTTGAACCTCTG 1859
|||||
Db 163 AGTGGATGGAGTGTCTGAGAGATCCTATGACAGCTGAGAGACCTGCGCTTGAACCTCTG 104
QY 1860 CCTGCTCCACACTTATCTCTTGAATATATGGGATGAGGTGATGAGGAAAGTTGGGG 1919
|||||
Db 103 CCTGCTCCACACTTATCTCTTGAATATATGGGATGAGGTGATGAGGAAAGTTGGGG 44
QY 1920 AAGTTTCTGTGTAAATAAAGGATCTTTCTTCAAAAA 1962
|||||
Db 43 AAGTTTCTGTGTAAATAAAGGATCTTTCTTCAAAAA 1

RESULT 35
AM823060/c 391 bp mRNA linear EST 17-MAY-2000
LOCUS ur19c12.x1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3156502 3',
DEFINITION mRNA sequence.
ACCESSION AM823060
VERSION AM823060.1 GI:7916137
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 391)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other_ESTs: ur19c12.y1
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
image.llnl.gov/image/html/iresources.shtml

MG1:1059258
Seq primer: -400p from Glibco
High quality sequence stop: 380.
FEATURES
location/Qualifiers
1..391
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3156502"
/clone_lib="NCI CGAP Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"

/note="Organ: mammary; Vector: PCMV-SPORT6; Site_1: Sal1;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT 107 a 106 c 87 g 90 t 1 others
ORIGIN

Query Match 19.7%; Score 388.4; DB 9; Length 391;
Best Local Similarity 99.5%; Pred. No. 4.1e-70;
Matches 389; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1563 ACCTGCTTAACCGAAGTGTGTTTGAAGATGAGTCTCTTGGACAGTCCAAAGAGAT 1622
|||||
Db 391 ACCTGCTTAACCGAAGTGTGTTTGAAGATGAGTCTCTTGGACAGTCCAAAGAGAT 332
QY 1623 GCCCAGAAAACACACTCTCTGTTCACATGCGCCCTGCACACACTGGGAAAGCCACATGA 1682
|||||
Db 331 GCCCAGAAAACACACTCTCTGTTCACATGCGCCCTGCACACACTGGGAAAGCCACATGA 272
QY 1683 CCAGTTTACTGTTCGATCAGACAGGCGCTACTCTCAGTTGCAAGGTTTGGCTTATACCTA 1742
|||||
Db 271 CCAGTTTACTGTTCGATCAGACAGGCGCTACTCTCAGTTGCAAGGTTTGGCTTATACCTA 212
QY 1743 CAACCAAGTGTGCTGAGACTCTTGTGTTTATAGAAGAGGTACATTCAGCTTAAGT 1802
|||||
Db 211 CAACCAAGTGTGCTGAGACTCTTGTGTTTATAGAAGAGGTACATTCAGCTTAAGT 152
QY 1803 GGATGGAGTGTCTGAGAGATCCTATGACAGCTGAGAGACCTGCGCTTGAACCTCTGCT 1862
|||||
Db 151 GGATGGAGTGTCTGAGAGATCCTATGACAGCTGAGAGACCTGCGCTTGAACCTCTGCT 92
QY 1863 GCGTCACACTTATGCTTGAATATATGGGATGAGGTGATGAGGAAAGTTGGGAG 1922
|||||
Db 91 GCGTCACACTTATGCTTGAATATATGGGATGAGGTGATGAGGAAAGTTGGGAG 32
QY 1923 TTTTCTGTGTAAATAAAGGATCTTTTC 1953
|||||
Db 31 TTTTCTGTGTAAATAAAGGATCTTTTC 1

RESULT 36
BB837528 393 bp mRNA linear EST 19-NOV-2001
LOCUS BB837528 RIKEN full-length enriched, mammary gland RCB-0527
DEFINITION Y9g-MC(B) cDNA Mus musculus cDNA clone G930047020 3', mRNA
sequence.
ACCESSION BB837528
VERSION BB837528.1 GI:17015771
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 393)
AUTHORS Akimura,T., Araiawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hiroaka,T., Hirozane,T., Imotani,K., Ishii
Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Koude,M., Matsuyama,T.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toyata,T.,
Watabiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)

JOURNAL Unpublished (2001)
COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,
URL: http://genome-gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
Wegli, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Matsubara, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Mitsuura
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multichannel sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome-gsc.riken.go.jp) for
further details.
e mouse tissues.

FEATURES

source

Location/Qualifiers
1. 393
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="G93004.7020"
/clone_lib="RIKEN full-length enriched, mammary gland
RCB-0527 JyG-MC(B) cDNA"
/tissue_type="mammary gland"
/cell_line="RCB-0527 JyG-MC(B)"
BASE COUNT 88 a 87 c 106 g 112 t
ORIGIN

Query Match 19.5%; Score 385.6; DB 9; Length 393;
Best Local Similarity 99.0%; Pred. No. 1.6e-69;
Matches 388; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1555 CTGGTACCGGAATGTTTGGAGAATGGCTCTCTTGGACCGAGTCCAGAGAGATGC 1624
Db 1 CTGGCTAACCGGAATGTTTGGAGAATGGCTCTCTTGGACCGAGTCCAGAGAGATGC 60
Oy 1625 CCAGAAACACACTCTCTGTGTTCACCTGCGCCCTGCACACACCTGGAGCCACATGACC 1684
Db 61 CCAGAAACACACTCTCTGTGTTCACCTGCGCCCTGCACACACCTGGAGCCACATGACC 120
Oy 1685 AGTTTACTGTTCCGATCAGACAGGGCCCTACTTCCAGTTGCAGGGTTTGTCTATAGCTACA 1744
Db 121 AGTTTACTGTTCCGATCAGACAGGGCCCTACTTCCAGTTGCAGGGTTTGTCTATAGCTACA 180
Oy 1745 ACCAGGTGTGGCTGAGACCTCTTTTGTATAGACAGGGTGCACATTGACTCTAAGTGG 1804
Db 181 ACCAGGTGTGGCTGAGACCTCTTTTGTATAGACAGGGTGCACATTGACTCTAAGTGG 240
Oy 1805 ATGGAGTGTGGAGATCTATGACAGCTGGAGGACCTGGGCTTGAACCTCTGCTGC 1864
Db 241 ATGGAGTGTGGAGATCTATGACAGCTGGAGGACCTGGGCTTGAACCTCTGCTGC 300
Oy 1865 CTCGAGCTTATGCTTGAATATGAGGTGAGTGTGATAGGAAAGTTGGGGAAGTT 1924
Db 301 CTCGAGCTTATGCTTGAATATGAGGTGAGTGTGATAGGAAAGTTGGGGAAGTT 360
Oy 1925 TTCTGTGTAATAAATAAAGGATCTTTCTTC 1956
Db 361 TTCTGTGTAATAAATAAAGGATCTTTCTTC 392

RESULT 37
BI341090 571 bp mRNA linear EST 30-JUL-2001
LOCUS BI341090
DEFINITION 368355 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BI341090

VERSION BI341090.1 GI:15034379
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
1 (bases 1 to 571)
Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.,
and Keele, J.W.

TITLE Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine

JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@meat.ars.usda.gov
Single pass sequencing. Bases called and alt. trimmed with phred
v0.980904.e. Vector identified by cross-match with the -minscore 18
and -mismatch 12 options.
PCR primers

FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 107 row: K column: 10
Seq primer: ATTTAGCTGACCTATGAC.

Location/Qualifiers

FEATURES

source

1. 571
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; site_1: XbaI; site_2: XhoI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 154 a 162 c 146 g 109 t
ORIGIN

Query Match 18.9%; Score 373.6; DB 10; Length 571;
Best Local Similarity 84.8%; Pred. No. 4.5e-67;
Matches 431; Conservative 0; Mismatches 74; Indels 3; Gaps 1;

Oy 81 TGTTCCTCGGCTGCTGAGTCAGACCATATGCTATCTCTGTGCTACTATCTG 140
Db 66 TGGTACCTGGGCGACCTGAGACAG--TCATGCTATTTGCTGCTGCTGCTACATATG 122
Oy 141 CTCCGACTTTCGATACATCCCGTGCAGCTGGCCATCCAGTGGGCCACTTTTCA 200
Db 123 CTCCGACTTTCGATACATCCCGTGCAGCTGGCCATCCAGTGGGCCACTTTTCA 182
Oy 201 TCTCAATGCTTAATTCAGGTTTGAAGACAGCAAGTGGAGCTGCCACAGTGTAG 260
Db 183 CTTCGAGTCTTAATTCAGGTTTGAAGACAGCAAGTGGAGCTGCCACAGTGTAG 242
Oy 261 AATCCAGTTGGCAAAAGACTATTAACAACACTTTTCTTGAACCTGCGCCAGAGA 320
Db 243 AATCCAGTTGGCAAAAGACTATTAACAACACTTTTCTTGAACCTGCGCCAGAGA 302
Oy 321 GGAGAATCTCTGATGACAGATTTTAAGAATGAAGTGCAGACGCTCAAGTCAAGCT 380
Db 303 GGAGAATCTCTGATGACAGATTTTAAGAATGAAGTGCAGATTAATACCAAGCTGCT 362
Oy 381 TTCCAGAAAGACAGAGAAAGGAGACAGGACCATTTATCGACACTAGCGGACAC 440
Db 363 TTCCAGAAAGACAGAGAAAGGAGACAGGACCATTTATCATCTATTGACCTTTGGCGGACAC 422
Oy 441 CTGGAAGAACGCAATGCTACCTGAGTCCCTTACAGAACGCTTAAACAAGGAGAT 500
Db 423 GCTGGAAGAGCGCAACCTGAGTGAATCCCTGACAGAAGGCTTAAACAAGGCGGAAT 482

